(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 October 2001 (18.10.2001)

PCT

(10) International Publication Number
WO 01/77331 A1

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, 16/18, C12Q 1/68, G01N 33/53, A61K 38/17, 39/395, 48/00

(21) International Application Number: PCT/US01/11442

(22) International Filing Date: 6 April 2001 (06.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/544,797 7 April 2000 (07.04.2000) US

(71) Applicant *(for all designated States except US)*: **MILLENNIUM PHARMACEUTICALS, INC.** [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CURTIS, Rory, A., J.** [GB/US]; 31 Constitution Drive, Southborough, MA 01772 (US). **SILOS-SANTIAGO, Inmaculada** [ES/US]; 18 Hilliard Street, Cambridge, MA 02138 (US).

(74) Agents: **MANDRAGOURAS, Amy, E.**; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: HUMAN TRP-LIKE CALCIUM CHANNEL PROTEIN-2 (TLCC-2)

Transmembrane Segments Predicted by MEMSAT

Start	End	Orient	Score
70	86	ins->out	2.0
299	317	out->ins	3.7
354	371	ins->out	2.4
385	402	out->ins	3.1
428	447	ins->out	4.7
497	521	out->ins	6.4

>54420
MTAPAGPRGSETERLLTPNPGYGTQAGPSPAPPTPPEEEDLRRRLKYFFMSPCDKFRAKG
RKPCKLMLQVVKILVVTVQLILFGLSNQLAVTFREENTIAFRHLFLLGYSDGADDTPAAY
TREQLYQAIFHAVDQYLALPDVSLGRYAYVRGGGDPWTNGSGLALCQRYYHRGHVDPAND
TPDIDPMVVTDCIQVDPPERPPPPPSDDLTLLESSSSYTNLTLKFHKLVNVTIHFRLKTI
NLQSLINNEIPDCYTFSVLITPDNKAHSGRIPISLETQAHIQECKHPSVPQHGDNSFRLL
FDVVVILTCSLSFLLCARSLLRGPLLQNEFVGFHWRQRGRVISLWERLEFVNGWYILLVT
SDVLTISGTIMKIGIEAKNLASYDVCSILLGTSTLLVMVGVIRYLTPFHNYNILIATLRV
ALPSVMRFCCCVAVIYLGYCFCGWIVLGPYHVKFRSLSMVSECLFSLINGDDDCVTFAAM
QAQQGRSSLVWLFSQLYLVSFISLFIYHVLSLFIALITGAYDTIKHPGGAGAEESELQAY
IAQCQDSPTSGKFRRGSGSACSLLCCCGRDPSEEHSLLVN

(57) Abstract: The invention provides isolated nucleic acids molecules, designated TLCC-2 nucleic acid molecules, which encode novel TRP-like calcium channel molecules. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing TLCC-2 nucleic acid molecules, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a TLCC-2 gene has been introduced or disrupted. The invention still further provides isolated TLCC-2 proteins, fusion proteins, antigenic peptides and anti-TLCC-2 antibodies. Diagnostic methods utilizing compositions of the invention are also provided.

WO 01/77331 A1

HUMAN TRP-LIKE CALCIUM CHANNEL PROTEIN-2 (TLCC-2)

Related Applications

The present application is a continuation-in-part of U.S. Patent Application Serial No. 09/544,797 entitled "54420, A NOVEL HUMAN CALCIUM CHANNEL", filed April 7, 2000. The content of the above-referenced patent application is incorporated herein by this reference in its entirety.

Background of the Invention

Calcium signaling has been implicated in the regulation of a variety of cellular responses, such as growth and differentiation. There are two general methods by which intracellular concentrations of calcium ions may be increased: calcium ions may be brought into the cell from the extracellular milieu through the use of specific channels in the cellular membrane, or calcium ions may be freed from intracellular stores, again being transported by specific membrane channels in the storage organelle. In the situation in which the intracellular stores of calcium have been depleted, a specific type of calcium channel, termed a 'capacitative calcium channel' or a 'store-operated calcium channel' (SOC), is activated in the plasma membrane to import calcium ions from the extracellular environment to the cytosol (for review, see Putney and McKay (1999) *BioEssays* 21:38-46).

Members of the capacitative calcium channel family include the calcium release-activated calcium current (CRAC) (Hoth and Penner (1992) *Nature* 355: 353-355), calcium release-activated nonselective cation current (CRANC) (Krause *et al.* (1996) *J. Biol. Chem.* 271: 32523-32528), and the transient receptor potential (TRP) proteins. There is no single electrophysological profile characteristic of the family; rather, a wide array of single channel conductances, cation selectivity, and current properties have been observed for different specific channels. Further, in several instances it has been demonstrated that homo- or heteropolymerization of the channel molecule may occur, further changing the channel properties from that of the single molecule. In general, though, these channels function similarly, in that they are calcium ion-permeable cation channels which become activated upon stimulation of phospholipase C_β by a G protein-coupled receptor. Depletion of intracellular calcium stores activate these channels by a

mechanism which is yet undefined, but which has been demonstrated to involve a diffusible factor using studies in which calcium stores were artificially depleted (*e.g.*, by the introduction of chelators into the cell, by activating phospholipase C_γ, or by inhibiting those enzymes responsible for pumping calcium ions into the stores or those enzymes responsible for maintaining resting intracellular calcium ion concentrations) (Putney, J.W., (1986) *Cell Calcium* 7: 1-12; Putney, J.W. (1990) *Cell Calcium* 11:611-624).

The TRP channel family is one of the best characterized members of the capacitative calcium channel group. These channels include transient receptor potential protein and homologues thereof (to date, seven homologs and splice variants have been identified in a variety of organisms), the vanilloid receptors (also known as the capsaicin receptors), stretch-inhibitable non-selective cation channel (SIC), olfactory, mechanosensitive channel, insulin-like growth factor I-regulated calcium channel, and vitamin D-responsive apical, epithelial calcium channel (ECaC), melastatin, and the polycystic kidney disease protein family (see, *e.g.*, Montell and Rubin (1989) *Neuron* 2:1313-1323; Caterina *et al.* (1997) *Nature* 389: 816-824; Suzuki *et al.* (1999) *J. Biol. Chem.* 274: 6330-6335; Kiselyov *et al.* (1998) *Nature* 396: 478-482; Hoenderop *et al.* (1999) *J. Biol. Chem.* 274: 8375-8378; and Chen et al. (1999) Nature 401(6751): 383-6). Each of these molecules is 700 or more amino acids in length (TRP and TRP homologs have 1300 or more amino acid residues), and shares certain conserved structural features. Predominant among these structural features are six transmembrane domains, with an additional hydrophobic loop present between the fifth and sixth transmembrane domains. It is believed that this loop is integral to the activity of the pore of the channel formed upon membrane insertion (Hardie and Minke (1993) *Trends Neurosci* 16: 371-376). TRP channel proteins also include one or more ankyrin domains and frequently display a proline-rich region at the N-terminus. Although found in disparate tissues and organisms, members of the TRP channel protein family all serve to transduce signals by means of calcium entry into cells, particularly pain signals (see, *e.g.*, McClesky and Gold (1999) *Annu. Rev. Physiol.* 61: 835-856), light signals (Hardie and Minke, *supra*), or olfactory signals (Colbert *et al.* (1997) *J. Neurosci* 17(21): 8259-8269). Thus, this family of molecules may play important roles in sensory signal transduction in general.

Vanilloid receptors (VRs) are non-selective cation channels that are structurally related to members of the TRP family of ion channels. These receptors have been proposed to mediate the entry of extracellular calcium into cells in response to the depletion of intracellular calcium stores. VRs are expressed in nociceptive neurons, as well as other cells types, and are activated by a variety of stimuli including noxious heat and protons. Capsaicin, which is a well-known agonist of VRs, induces pain behavior in humans and rodents. VR-1, a vanilloid receptor, was identified in rat sensory ganglia (Caterina M. J. et al., (1997) *Nature* 389:816-824). It has been shown that VR-1 knockout mice are impaired in their detection of painful heat, exhibit no vanilloid-evoked pain behavior, and show little thermal hypersensitivity after inflammation (Szallasi and Blumberg (1999) Pharmacol. Rev. 51:159-211; Tominaga, *et al.* (1998) *Neuron* 21:531; Caterina *et al.* (2000) *Science* 288:306).

Summary of the Invention

The present invention is based, at least in part, on the discovery of novel transient receptor potential (TRP) (*e.g.,* the calcium channel and/or vanilloid receptor) family members, referred to herein as TRP-like calcium channel or TLCC-2 nucleic acid and protein molecules. The TLCC-2 molecules of the present invention are useful as targets for developing modulating agents to regulate a variety of cellular processes, including membrane excitability, neurite outgrowth and synaptogenesis, signal transduction, cell proliferation, growth, differentiation, and migration, and nociception. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding TLCC-2 proteins or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of TLCC-2-encoding nucleic acids.

In one embodiment, a TLCC-2 nucleic acid molecule of the invention is at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the nucleotide sequence (*e.g.*, to the entire length of the nucleotide sequence) shown in SEQ ID NO:1 or 3 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a complement thereof.

In a preferred embodiment, the isolated nucleic acid molecule includes the nucleotide sequence shown SEQ ID NO:1 or 3, or a complement thereof. In another embodiment, the nucleic acid molecule includes SEQ ID NO:3 and nucleotides 1-140 of SEQ ID NO:1. In yet another embodiment, the nucleic acid molecule includes SEQ ID NO:3 and nucleotides 1884-2095 of SEQ ID NO:1. In another preferred embodiment, the nucleic acid molecule consists of the nucleotide sequence shown in SEQ ID NO:1 or 3. In another preferred embodiment, the nucleic acid molecule includes a fragment of at least 706 nucleotides (*e.g.*, 706 contiguous nucleotides) of the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, or a complement thereof.

In another embodiment, a TLCC-2 nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In a preferred embodiment, a TLCC-2 nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the entire length of the amino acid sequence of SEQ ID NO:2 or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

In another preferred embodiment, an isolated nucleic acid molecule encodes the amino acid sequence of human TLCC-2. In yet another preferred embodiment, the nucleic acid molecule includes a nucleotide sequence encoding a protein having the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, the nucleic acid molecule is at least 706 nucleotides in length. In a further preferred embodiment, the nucleic acid molecule is at least 706 nucleotides in length and encodes a protein having a TLCC-2 activity (as described herein).

Another embodiment of the invention features nucleic acid molecules, preferably TLCC-2 nucleic acid molecules, which specifically detect TLCC-2 nucleic acid molecules relative to nucleic acid molecules encoding non-TLCC-2 proteins. For example, in one embodiment, such a nucleic acid molecule is at least 706, 706-750, 750-800, 800-850, 850-900, 900-950, 950-1000, 1000-1050, 1050-1070, 1070-1100, 1100-

1150, 1150-1200, 1200-1250, 1250-1300, 1300-1350, 1350-1400, 1400-1450, 1450-1500, 1500-1550, 1550-1600, 1600-1650, 1650-1700, 1700-1750, 1750-1800, 1800-1850, 1850-1900, 1900-1950, 1950-2000, 2000-2050, 2050-2100 or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence shown in SEQ ID NO:1, the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number_____, or a complement thereof.

In preferred embodiments, the nucleic acid molecules are at least 15 (*e.g.*, 15 contiguous) nucleotides in length and hybridize under stringent conditions to nucleotides 1-27 of SEQ ID NO:1.

In other preferred embodiments, the nucleic acid molecule encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:1 or 3 under stringent conditions.

Another embodiment of the invention provides an isolated nucleic acid molecule which is antisense to a TLCC-2 nucleic acid molecule, *e.g.*, the coding strand of a TLCC-2 nucleic acid molecule.

Another aspect of the invention provides a vector comprising a TLCC-2 nucleic acid molecule. In certain embodiments, the vector is a recombinant expression vector. In another embodiment, the invention provides a host cell containing a vector of the invention. In yet another embodiment, the invention provides a host cell containing a nucleic acid molecule of the invention. The invention also provides a method for producing a protein, preferably a TLCC-2 protein, by culturing in a suitable medium, a host cell, *e.g.*, a mammalian host cell such as a non-human mammalian cell, of the invention containing a recombinant expression vector, such that the protein is produced.

Another aspect of this invention features isolated or recombinant TLCC-2 proteins and polypeptides. In one embodiment, an isolated TLCC-2 protein has one or more of the following domains: a transmembrane domain, a pore domain, and a proline rich domain. In a preferred embodiment, a TLCC-2 protein includes at least one or more of the following domains: a transmembrane domain, a pore domain, and a proline

rich domain, and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another preferred embodiment, a TLCC-2 protein includes at least one transmembrane domain and has a TLCC-2 activity (as described herein).

In yet another preferred embodiment, a TLCC-2 protein includes one or more of the following domains: a transmembrane domain, a pore domain, and a proline rich domain, and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3.

In another embodiment, the invention features fragments of the protein having the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 18 or more amino acids (*e.g.*, contiguous amino acids) of the amino acid sequence of SEQ ID NO:2, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number_____. In another embodiment, a TLCC-2 protein has the amino acid sequence of SEQ ID NO:2.

In another embodiment, the invention features a TLCC-2 protein which is encoded by a nucleic acid molecule consisting of a nucleotide sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to a nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof. This invention further features a TLCC-2 protein, which is encoded by a nucleic acid molecule consisting of a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof.

The proteins of the present invention or portions thereof, *e.g.*, biologically active portions thereof, can be operatively linked to a non-TLCC-2 polypeptide (*e.g.*, heterologous amino acid sequences) to form fusion proteins. The invention further features antibodies, such as monoclonal or polyclonal antibodies, that specifically bind proteins of the invention, preferably TLCC-2 proteins. In addition, the TLCC-2 proteins or biologically active portions thereof can be incorporated into pharmaceutical compositions, which optionally include pharmaceutically acceptable carriers.

In another aspect, the present invention provides a method for detecting the presence of a TLCC-2 nucleic acid molecule, protein, or polypeptide in a biological sample by contacting the biological sample with an agent capable of detecting a TLCC-2 nucleic acid molecule, protein, or polypeptide such that the presence of a TLCC-2 nucleic acid molecule, protein or polypeptide is detected in the biological sample.

In another aspect, the present invention provides a method for detecting the presence of TLCC-2 activity in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of TLCC-2 activity such that the presence of TLCC-2 activity is detected in the biological sample.

In another aspect, the invention provides a method for modulating TLCC-2 activity comprising contacting a cell capable of expressing TLCC-2 with an agent that modulates TLCC-2 activity such that TLCC-2 activity in the cell is modulated. In one embodiment, the agent inhibits TLCC-2 activity. In another embodiment, the agent stimulates TLCC-2 activity. In one embodiment, the agent is an antibody that specifically binds to a TLCC-2 protein. In another embodiment, the agent modulates expression of TLCC-2 by modulating transcription of a TLCC-2 gene or translation of a TLCC-2 mRNA. In yet another embodiment, the agent is a nucleic acid molecule having a nucleotide sequence that is antisense to the coding strand of a TLCC-2 mRNA or a TLCC-2 gene.

In one embodiment, the methods of the present invention are used to treat a subject having a disorder characterized by aberrant or unwanted TLCC-2 protein or nucleic acid expression or activity, *e.g.*, a pain disorder, by administering an agent which is a TLCC-2 modulator to the subject. In one embodiment, the TLCC-2 modulator is a TLCC-2 protein. In another embodiment the TLCC-2 modulator is a TLCC-2 nucleic acid molecule. In yet another embodiment, the TLCC-2 modulator is an antibody, ribozyme, peptide, peptidomimetic, antisense oligonucleotide, or other small molecule. In a preferred embodiment, the disorder characterized by aberrant or unwanted TLCC-2 protein or nucleic acid expression is a CNS disorder, such as a neurodegenerative disorder, *e.g.*, Alzheimer's disease, dementias related to Alzheimer's disease (such as Pick's disease), Parkinson's and other Lewy diffuse body diseases, multiple sclerosis, amyotrophic lateral sclerosis, progressive supranuclear palsy, epilepsy, Jakob-Creutzfieldt disease, or AIDS related dementia, familial infantile

convulsions, paroxysmal choreoathetosis; a disorder of the conveyance of sensory impulses from the periphery to the brain and/or conductance of motor impulses from the brain to the periphery; a psychiatric disorder (*e.g.*, depression, schizophrenic disorders, korsakoff's psychosis, mania, anxiety disorders, or phobic disorders; or a learning or memory disorder (*e.g.*, amnesia or age-related memory loss; or is a neurological disorder (*e.g.*, migraine).

In another embodiment, the disorder characterized by aberrant or unwanted TLCC-2 activity is a pain disorder, or a disorder characterized by misregulated pain signaling mechanisms.

In another embodiment, the disorder characterized by aberrant or unwanted TLCC-2 activity is a cell proliferation, growth, differentiation, or migration disorder.

The present invention also provides diagnostic assays for identifying the presence or absence of a genetic alteration characterized by at least one of (i) aberrant modification or mutation of a gene encoding a TLCC-2 protein; (ii) mis-regulation of the gene; and (iii) aberrant post-translational modification of a TLCC-2 protein, wherein a wild-type form of the gene encodes a protein with a TLCC-2 activity.

In another aspect the invention provides methods for identifying a compound that binds to or modulates the activity of a TLCC-2 protein, by providing an indicator composition comprising a TLCC-2 protein having TLCC-2 activity, contacting the indicator composition with a test compound, and determining the effect of the test compound on TLCC-2 activity in the indicator composition to identify a compound that modulates the activity of a TLCC-2 protein.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Drawings

Figure 1A-B depicts the cDNA sequence and predicted amino acid sequence of human TLCC-2. The nucleotide sequence corresponds to nucleic acids 1 to 2095 of SEQ ID NO:1. The amino acid sequence corresponds to amino acids 1 to 580 of SEQ ID NO:2. The coding region without the 3' untranslated region of the human TLCC-2 gene is shown in SEQ ID NO:3.

Figure 2 depicts a structural, hydrophobicity, and antigenicity analysis of the human TLCC-2 protein.

Figure 3 depicts the results of a search which was performed against the MEMSAT database and which resulted in the identification of six "transmembrane domains" in the human TLCC-2 protein (SEQ ID NO:2).

Figure 4 depicts the results of a search which was performed against the HMM database and which resulted in the identification of a "Fibronectin type III domain" in the human TLCC-2 protein (SEQ ID NO:2).

Figure 5 depicts the results of a search which was performed against the ProDom database and which resulted in the identification of a "31K RNA-4 protein domain" in the human TLCC-2 protein (SEQ ID NO:2).

Figure 6 depicts an alignment of the human TLCC-2 amino acid sequence (SEQ ID NO:2) with the amino acid sequences of human vanilloid receptor 1 and human vanilloid receptor 2 (Accession Numbers AAD26363 and AAD41724, set forth as SEQ ID NO:4 and SEQ ID NO:5, respectively), using the CLUSTAL W(1.74) alignment program.

Figure 7A-B depicts an alignment of the human TLCC-2 amino acid sequence (SEQ ID NO:2) with the amino acid sequence of polycystic kidney disease protein 2 from *Mus musculus* (Accession Number NP_032887, set forth as SEQ ID NO:6), using the CLUSTAL W(1.74) alignment program.

Figure 8A-B depicts an alignment of the human TLCC-2 amino acid sequence (SEQ ID NO:2) with the amino acid sequence of human melastatin (Accession Number AAC8000, set forth as SEQ ID NO:7), using the CLUSTAL W(1.74) alignment program.

Figure 9 depicts the results of a search performed against the Prosite database and which resulted in the identification of four N-glycosylation sites in the amino acid sequence of human TLCC-2 (SEQ ID NO:2).

Figure 10 is a graphic depiction of the relative levels of the human TLCC-2 mRNA expression in a human normal tissue panel, as determined using Taqman™ analysis.

Figure 11 is a graphic depiction of the relative levels of the human TLCC-2 mRNA expression in a human normal tissue panel, as determined using Taqman™ analysis.

Detailed Description of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as " TRP-like calcium channel" or "TLCC-2" nucleic acid and protein molecules, which are novel members of the ion channel, *e.g.,* calcium channel and/or vanilloid receptor, family. These novel molecules are capable of, for example, modulating an ion-channel mediated activity (*e.g.*, a calcium channel- and/or vanilloid receptor-mediated activity) in a cell, *e.g.*, a neuronal, skin, muscle (*e.g.*, cardiac muscle), or liver cell.

The present invention is based also in part, on the discovery that the novel TLCC-2 molecules of the present invention are expressed in the brain at high levels, and are also expressed in the skin, the spinal cord and dorsal root ganglia (DRG). Moreover, the novel TLCC-2 molecules of the invention are upregulated in animal models of pain. The TLCC-2 molecules of the invention are involved in nociception (*e.g.*, chemical, mechanical, or thermal nociception) and thereby modulate pain elicitation. Accordingly, the TLCC-2 molecules of the present invention act as targets for developing novel diagnostic targets and therapeutic agents to control pain and pain disorders.

As used herein, an "ion channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting signals in an electrically excitable cell, *e.g.*, a neuronal or muscle cell. Ion channels include vanilloid receptors, calcium channels, potassium channels, and sodium channels. As used herein, a "calcium channel" includes a protein or polypeptide which is involved in receiving, conducting, and trasmitting calcium ion-based signals in an electrically excitable cell. Calcium channels are calcium ion selective, and can determine membrane excitability (the ability of, for example, a neuronal cell to respond to a stimulus and to convert it into a sensory impulse). Calcium channels can also influence the resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation. Calcium channels are typically expressed in electrically excitable cells, *e.g.*, neuronal cells, and may form heteromultimeric structures (*e.g.*, composed of more than one type of

subunit). Calcium channels may also be found in non-excitable cells (*e.g.*, adipose cells or liver cells), where they may play a role in, *e.g.*, signal transduction.. Examples of calcium channels include the low-voltage-gated channels and the high-voltage-gated channels. Calcium channels are described in, for example, Davila *et al.* (1999) *Annals New York Academy of Sciences* 868:102-17 and McEnery, M.W. *et al.* (1998) *J. Bioenergetics and Biomembranes* 30(4): 409-418, the contents of which are incorporated herein by reference.

As used herein, a "vanilloid receptor" includes a non- selective cation channel that is structurally related to the TRP family of ion channels. Vanilloid receptors are also known as capsaisin receptors. Vanilloid receptors share several physical characteristics including an N-terminal cytoplasmic domain which contains three ankyrin repeats, six transmembrane domains, a pore-loop region located between transmembrane domains 5 and 6, and several kinase consensus sequences. Members of the vanilloid receptor (VR) family have been proposed to mediate the entry of extracellular calcium into cells, *e.g.*, in response to the depletion of intracellular calcium stores. VRs are typically expressed in nociceptive neurons among other cells types and are directly activated by harmful heat, extracellular protons, and vanilloid compounds. VRs may also be expressed in nonsensory tissues and may mediate inflammatory rather than acute thermal pain. Vanilloid receptors are described in, for example, Caterina, M.J. (1997) *Nature* 389:816-824 and Caterina, M.J. (2000) *Science* 288:306-313) the contents of which are incorporated herein by reference. As the TLCC-2 molecules of the present invention may modulate ion channel mediated activities (*e.g.*, calcium channel- and/or vanilloid receptor- mediated activities), they may be useful for developing novel diagnostic and therapeutic agents for ion channel associated disorders (*e.g.*, calcium channel and/or vanilloid receptor associated disorders).

As used herein, an "ion channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of ion channel (*e.g.*, calcium channel) and/or vanilloid receptor) mediated activity. For example, a "calcium channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of calcium channel mediated activity. A "vanilloid receptor associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of vanilloid receptor mediated activity. Ion channel associated disorders,

e.g., calcium channel associated disorders and/or vanilloid receptor associated disorders, include CNS disorders, such as cognitive and neurodegenerative disorders, examples of which include, but are not limited to, Alzheimer's disease, dementias related to Alzheimer's disease (such as Pick's disease), Parkinson's and other Lewy diffuse body diseases, senile dementia, Huntington's disease, Gilles de la Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis, progressive supranuclear palsy, epilepsy, Jakob-Creutzfieldt disease, or AIDS related dementia; autonomic function disorders such as hypertension and sleep disorders, and neuropsychiatric disorders, such as depression , schizophrenia, schizoaffective disorder, korsakoff's psychosis, mania, anxiety disorders, or phobic disorders; leaning or memory disorders, *e.g.*, amnesia or age-related memory loss, attention deficit disorder, psychoactive substance use disorders, anxiety, phobias, panic disorder, as well as bipolar affective disorder, *e.g.*, severe bipolar affective (mood) disorder (BP-1), and bipolar affective neurological disorders, *e.g.*, migraine and obesity. Further CNS-related disorders include, for example, those listed in the American Psychiatric Association's Diagnostic and Statistical manual of Mental Disorders (DSM), the most current version of which is incorporated herein by reference in its entirety.

Ion channel associated disorders, *e.g.*, calcium channel disorders and/or vanilloid receptor associated disorders, also include pain disorders. As used herein, the term "pain disorder" includes a disorder affecting pain signaling mechanisms. Pain disorders include disorders characterized by aberrant (*e.g.*, excessive or amplified) pain. Examples of pain disorders include posttherapeutic neuralgia, diabetic neuropathy, postmastectomy pain syndrome, stump pain, reflex sympathetic dystrophy, trigeminal neuralgia, neuropathic pain, orofacial neuropathic pain, osteoarthritis, rheumatoid arthritis, fibromyalgia syndrome, tension myalgia, Guillian-Barre syndrome, Meralgia paraesthetica, burning mouth syndrome, fibrocitis, myofascial pain syndrome, idiopathic pain disorder, temporomandibular joint syndrome, atypical odontalgia, loin pain, haematuria syndrome, non-cardiac chest pain, low back pain, chronic nonspecific pain, psychogenic pain, musculoskeletal pain disorder, chronic pelvic pain, nonorganic chronic headache, tension-type headache, cluster headache, migraine, complex regional pain syndrome, vaginismus, nerve trunk pain, somatoform pain disorder, cyclical mastalgia, chronic fatigue syndrome, multiple somatization syndrome, chronic pain

disorder, somatization disorder, Syndrome X, facial pain, idiopathic pain disorder, posttraumatic rheumatic pain modulation disorder (fibrositis syndrome), hyperalgesia, and Tangier disease.

As used herein, the term "pain signaling mechanisms" includes the cellular mechanisms involved in the development and regulation of pain, *e.g.*, pain elicited by noxious chemical, mechanical, or thermal stimuli, in a subject, *e.g.*, a mammal such as a human. In mammals, the initial detection of noxious chemical, mechanical, or thermal stimuli, a process referred to as "nociception", occurs predominantly at the peripheral terminals of specialized, small diameter sensory neurons. These sensory neurons transmit the information to the central nervous system, evoking a perception of pain or discomfort and initiating appropriate protective reflexes. The TLCC-2 molecules of the present invention may be present on these sensory neurons and, thus, may be involved in detecting these noxious chemical, mechanical, or thermal stimuli and transducing this information into membrane depolarization events. Thus, the TLCC-2 molecules, by participating in pain signaling mechanisms, may modulate pain elicitation and act as targets for developing novel diagnostic targets and therapeutic agents to control pain.

Ion channel associated disorders, *e.g.*, calcium channel disorders and/or vanilloid receptor associated disorders, also include cellular proliferation, growth, differentiation, or migration disorders. Cellular proliferation, growth, differentiation, or migration disorders include those disorders that affect cell proliferation, growth, differentiation, or migration processes. As used herein, a "cellular proliferation, growth, differentiation, or migration process" is a process by which a cell increases in number, size or content, by which a cell develops a specialized set of characteristics which differ from that of other cells, or by which a cell moves closer to or further from a particular location or stimulus. The TLCC-2 molecules of the present invention are involved in signal transduction mechanisms, which are known to be involved in cellular growth, differentiation, and migration processes. Thus, the TLCC-2 molecules may modulate cellular growth, differentiation, or migration, and may play a role in disorders characterized by aberrantly regulated growth, differentiation, or migration. Such disorders include cancer, *e.g.*, carcinoma, sarcoma, or leukemia; tumor angiogenesis and metastasis; skeletal dysplasia; neuronal deficiencies resulting from impaired neural induction and

patterning; hepatic disorders; cardiovascular disorders; and hematopoietic and/or myeloproliferative disorders.

As used herein, an "ion channel mediated activity" includes an activity which involves an ion channel, *e.g.*, an ion channel in a neuronal cell, a skin cell, a muscular cell, or a liver cell, associated with receiving, conducting, and transmitting signals, in, for example, the nervous system. Ion channel mediated activities (*e.g.*, calcium channel and/or vanilloid receptor mediated activities) include release of neurotransmitters or second messenger molecules (*e.g.*, dopamine or norepinephrine), from cells, *e.g.*, neuronal cells; mediation of entry of extracellular calcium into cells, *e.g.*, neuronal cells; modulation of resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation; participation in signal transduction pathways, and modulation of processes such as integration of sub-threshold synaptic responses and the conductance of back-propagating action potentials in, for example, neuronal cells (*e.g.*, changes in those action potentials resulting in a morphological or differentiative response in the cell).

The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. Such family members can be naturally or non-naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of human origin or alternatively, can contain homologues of non-human origin, *e.g.*, monkey proteins. Members of a family may also have common functional characteristics.

For example, the family of TLCC-2 proteins comprise at least one "transmembrane domain" and preferably six transmembrane domains. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 20-45 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a

transmembrane domain are hydrophobic, *e.g.*, leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W.N. *et al,* (1996) *Annual Rev. Neurosci.* 19: 235-263, the contents of which are incorporated herein by reference. Amino acid residues 70-86, 299-317, 354-371, 385-416, 428-447, and 497-521 of the TLCC-2 protein comprise transmembrane domains (see Figures 2 and 3). Accordingly, TLCC-2 proteins having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a transmembrane domain of human TLCC-2 are within the scope of the invention.

In another embodiment, a TLCC-2 molecule of the present invention is identified based on the presence of at least one pore domain between the fifth and sixth transmembrane domains. As used herein, the term "pore domain" includes an overall hydrophobic amino acid sequence which is located between two transmembrane domains of a calcium channel protein, preferably transmembrane domains 5 and 6, and which is believed to be a major determinant of ion selectivity and channel activity in calcium channels. Pore domains are described, for example in Vannier *et al.* (1998) *J. Biol. Chem.* 273: 8675-8679 and Phillips, A. M. *et al.* (1992) *Neuron* 8, 631-642, the contents of which are incorporated herein by reference. TLCC-2 molecules having at least one pore domain are within the scope of the invention. A pore domain may be found in the human TLCC-2 sequence (SEQ ID NO:2) at about residues 459-470 (Figure 2).

In another embodiment, a TLCC-2 molecule of the present invention is identified based on the presence of at least one N-glycosylation site. As used herein, the term "N-glycosylation site" includes an amino acid sequence of about 4 amino acid residues in length which serves as a glycosylation site. More preferably, an N-glycosylation site has the consensus sequence Asn-Xaa-Ser/Thr (where Xaa may be any amino acid) (SEQ ID NO:4). N-glycosylation sites are described in, for example, Prosite PDOC00001 (http://www.expasy.ch/cgi-bin/get-prodoc-entry?PDOC00001), the contents of which are incorporated herein by reference. Amino acid residues 159-162, 179-182, 220-223, and 230-233 of the TLCC-2 protein comprise N-glycosylation sites (see Figure 9). Accordingly, TLCC-2 proteins having at least one N-glycosylation site are within the scope of the invention.

In another embodiment, a TLCC-2 molecules of the present invention is identified based on the presence of a "proline rich domain" in the protein or corresponding nucleic acid molecule. As used herein, the term "proline rich domain" includes an amino acid sequence of about 4-6 amino acid residues in length having the general sequence X-Pro-X-X-Pro-X (where X can be any amino acid). Proline rich domains are usually located in a helical structure and bind through hydrophobic interactions to SH3 domains. SH3 domains recognize proline rich domains in both forward and reverse orientations. Proline rich domains are described in, for example, Sattler, M. *et al.* (1998) *Leukemia* 12: 637-644, the contents of which are incorporated herein by reference. Residues 1-37 of the amino acid sequence of human TLCC-2 (SEQ ID NO:2) contain proline-rich domains.

In a preferred embodiment, the TLCC-2 molecules of the invention include at least one transmembrane domain, at least one N-glycosylation site, at least one pore domain, and at least one proline rich domain.

Isolated proteins of the present invention, preferably TLCC-2 proteins, have an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2 or are encoded by a nucleotide sequence sufficiently identical to SEQ ID NO:1 or 3. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (*e.g.*, an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains have at least 30%, 40%, or 50% homology, preferably 60% homology, more preferably 70%-80%, and even more preferably 90-95% homology across the amino acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently identical. Furthermore, amino acid or nucleotide sequences which share at least 30%, 40%, or 50%, preferably 60%, more preferably 70-80%, or 90-95% homology and share a common functional activity are defined herein as sufficiently identical.

As used interchangeably herein, an "TLCC-2 activity", "biological activity of TLCC-2 " or "functional activity of TLCC-2", refers to an activity exerted by a TLCC-2 protein, polypeptide or nucleic acid molecule on a TLCC-2 responsive cell or tissue, or on a TLCC-2 protein substrate, as determined *in vivo*, or *in vitro*, according to standard techniques. In one embodiment, a TLCC-2 activity is a direct activity, such as an association with a TLCC-2-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with which a TLCC-2 protein binds or interacts in nature, such that TLCC-2-mediated function is achieved. A TLCC-2 target molecule can be a non-TLCC-2 molecule or a TLCC-2 protein or polypeptide of the present invention: In an exemplary embodiment, a TLCC-2 target molecule is a TLCC-2 ligand, *e.g.*, a calcium channel ligand. Alternatively, a TLCC-2 activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the TLCC-2 protein with a TLCC-2 ligand. The biological activities of TLCC-2 are described herein. For example, the TLCC-2 proteins of the present invention can have one or more of the following activities: (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials, (4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction, and (7) participate in nociception.

Accordingly, another embodiment of the invention features isolated TLCC-2 proteins and polypeptides having a TLCC-2 activity. Preferred proteins are TLCC-2 proteins having at least one or more of the following domains: a transmembrane domain, an N-glycosylation site, a pore domain, and a proline rich domain, and, preferably, a TLCC-2 activity.

Additional preferred proteins have one or more of the following domains: a transmembrane domain, an N-glycosylation site, a pore domain, and a proline rich domain, and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3.

The nucleotide sequence of the isolated human TLCC-2 cDNA and the predicted amino acid sequence of the human TLCC-2 polypeptide are shown in Figure 1 and in SEQ ID NOs:1 and 2, respectively. A plasmid containing the nucleotide sequence encoding human TLCC-2 was deposited with the American Type Culture Collection

(ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on _____ and assigned Accession Number _____. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The human TLCC-2 gene, which is approximately 2095 nucleotides in length, encodes a protein having a molecular weight of approximately 65.7 kD and which is approximately 580 amino acid residues in length.

Various aspects of the invention are described in further detail in the following subsections:

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode TLCC-2 proteins or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify TLCC-2-encoding nucleic acid molecules (*e.g.*, TLCC-2 mRNA) and fragments for use as PCR primers for the amplification or mutation of TLCC-2 nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

The term "isolated nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated TLCC-2 nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA

of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, as a hybridization probe, TLCC-2 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to TLCC-2 nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO:1. The sequence of SEQ ID NO:1 corresponds to the human TLCC-2 cDNA. This cDNA comprises sequences encoding the human TLCC-2 protein (*i.e.*, "the coding region", from nucleotides 141-1883), as well as 5' untranslated sequences (nucleotides 1-140) and 3' untranslated

sequences (nucleotides 1884-2095). Alternatively, the nucleic acid molecule can comprise only the coding region of SEQ ID NO:1 (*e.g.*, nucleotides 141-1883, corresponding to SEQ ID NO:3).

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ , or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the entire length of the nucleotide sequence shown in SEQ ID NO:1 or 3, or the entire length of the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a TLCC-2 protein, *e.g.*, a biologically active portion of a TLCC-2 protein. The nucleotide sequence determined from the cloning of the TLCC-2 gene allows for the generation of probes and primers designed for use in identifying and/or cloning other TLCC-2 family members, as well as TLCC-2 homologues from other species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes

under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, 75, 80, 85, 90, 95, or 100 or more consecutive nucleotides of a sense sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, of an anti-sense sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In one embodiment, a nucleic acid molecule of the present invention comprises a nucleotide sequence which is greater than 706, 706-750, 750-800, 800-850, 850-900, 900-950, 950-1000, 1000-1050, 1050-1070, 1070-1100, 1100-1150, 1150-1200, 1200-1250, 1250-1300, 1300-1350, 1350-1400, 1400-1450, 1450-1500, 1500-1550, 1550-1600, 1600-1650, 1650-1700, 1700-1750, 1750-1800, 1800-1850, 1850-1900, 1900-1950, 1950-2000, 2000-2050, 2050-2100 or more nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

Probes based on the TLCC-2 nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a TLCC-2 protein, such as by measuring a level of a TLCC-2-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting TLCC-2 mRNA levels or determining whether a genomic TLCC-2 gene has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion of a TLCC-2 protein" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, which encodes a polypeptide having a TLCC-2 biological activity (the biological activities of the TLCC-2 proteins are described herein), expressing the encoded portion of the TLCC-2 protein (*e.g.*, by recombinant

expression *in vitro*) and assessing the activity of the encoded portion of the TLCC-2 protein.

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, due to degeneracy of the genetic code and thus encode the same TLCC-2 proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:2.

In addition to the TLCC-2 nucleotide sequences shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the TLCC-2 proteins may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the TLCC-2 genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a TLCC-2 protein, preferably a mammalian TLCC-2 protein, and can further include non-coding regulatory sequences, and introns.

Allelic variants of human TLCC-2 include both functional and non-functional TLCC-2 proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the human TLCC-2 protein that maintain the ability to bind a TLCC-2 ligand or substrate and/or modulate membrane excitability or signal transduction. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2, or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human TLCC-2 protein that do not have the ability to form functional calcium channels or to modulate membrane excitability. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion or

premature truncation of the amino acid sequence of SEQ ID NO:2, or a substitution, insertion or deletion in critical residues or critical regions.

The present invention further provides non-human orthologues of the human TLCC-2 proteins. Orthologues of the human TLCC-2 protein are proteins that are isolated from non- non-human organisms and possess the same TLCC-2 ligand binding and/or modulation of membrane excitation mechanisms of the human TLCC-2 protein. Orthologues of the human TLCC-2 protein can readily be identified as comprising an amino acid sequence that is substantially identical to SEQ ID NO:2.

Moreover, nucleic acid molecules encoding other TLCC-2 family members and, thus, which have a nucleotide sequence which differs from the TLCC-2 sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, another TLCC-2 cDNA can be identified based on the nucleotide sequence of human TLCC-2. Moreover, nucleic acid molecules encoding TLCC-2 proteins from different species, and which, thus, have a nucleotide sequence which differs from the TLCC-2 sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, a mouse TLCC-2 cDNA can be identified based on the nucleotide sequence of a human TLCC-2.

Nucleic acid molecules corresponding to natural allelic variants and homologues of the TLCC-2 cDNAs of the invention can be isolated based on their homology to the TLCC-2 nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Nucleic acid molecules corresponding to natural allelic variants and homologues of the TLCC-2 cDNAs of the invention can further be isolated by mapping to the same chromosome or locus as the TLCC-2 gene.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15, 20, 25, 30 or more nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In other embodiment, the nucleic acid is at least 706, 706-750, 750-800, 800-850, 850-900, 900-950, 950-1000, 1000-

1050, 1050-1070, 1070-1100, 1100-1150, 1150-1200, 1200-1250, 1250-1300, 1300-1350, 1350-1400, 1400-1450, 1450-1500, 1500-1550, 1550-1600, 1600-1650, 1650-1700, 1700-1750, 1750-1800, 1800-1850, 1850-1900, 1900-1950, 1950-2000, 2000-2050, 2050-2100 or more nucleotides in length.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45° C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length,

T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS), chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH_2PO_4, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH_2PO_4, 1% SDS at 65°C, see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).

Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1 or 3 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In addition to naturally-occurring allelic variants of the TLCC-2 sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby leading to changes in the amino acid sequence of the encoded TLCC-2 proteins, without altering the functional ability of the TLCC-2 proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of TLCC-2 (*e.g.*, the sequence of SEQ ID NO:2) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the TLCC-2 proteins of the present invention, *e.g.*, those present in a

transmembrane domain, are predicted to be particularly unamenable to alteration. Furthermore, additional amino acid residues that are conserved between the TLCC-2 proteins of the present invention and other members of the TLCC-2 family are not likely to be amenable to alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding TLCC-2 proteins that contain changes in amino acid residues that are not essential for activity. Such TLCC-2 proteins differ in amino acid sequence from SEQ ID NO:2, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:2.

An isolated nucleic acid molecule encoding a TLCC-2 protein identical to the protein of SEQ ID NO:2, can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a TLCC-2 protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be

introduced randomly along all or part of a TLCC-2 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for TLCC-2 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In a preferred embodiment, a mutant TLCC-2 protein can be assayed for the ability to: (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials, (4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction,(7) participate in nociception, and (8) modulate pain signaling mechanisms.

In addition to the nucleic acid molecules encoding TLCC-2 proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire TLCC-2 coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding TLCC-2. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the coding region of human TLCC-2 corresponds to SEQ ID NO:3). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding TLCC-2. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding TLCC-2 disclosed herein (*e.g.*, SEQ ID NO:3), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of TLCC-2 mRNA, but more preferably is an

oligonucleotide which is antisense to only a portion of the coding or noncoding region of TLCC-2 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of TLCC-2 mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a TLCC-2 protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by

conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids. Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave TLCC-2 mRNA transcripts to thereby inhibit translation of TLCC-2 mRNA. A ribozyme having specificity for a TLCC-2-encoding nucleic acid can be designed based upon the nucleotide sequence of a TLCC-2 cDNA disclosed herein (*i.e.*, SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a TLCC-2-

encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, TLCC-2 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, TLCC-2 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the TLCC-2 (*e.g.*, the TLCC-2 promoter and/or enhancers; *e.g.*, nucleotides 1-137 of SEQ ID NO:1) to form triple helical structures that prevent transcription of the TLCC-2 gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

In yet another embodiment, the TLCC-2 nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.* *Proc. Natl. Acad. Sci.* 93: 14670-675.

PNAs of TLCC-2 nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of TLCC-2 nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (*e.g.*, by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.*, S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

In another embodiment, PNAs of TLCC-2 can be modified, (*e.g.*, to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of TLCC-2 nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (*e.g.*, RNAse H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, *e.g.*, Krol *et al.* (1988) *Bio-Techniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (*e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

Alternatively, the expression characteristics of an endogenous TLCC-2 gene within a cell line or microorganism may be modified by inserting a heterologous DNA regulatory element into the genome of a stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous TLCC-2 gene. For example, an endogenous TLCC-2 gene which is normally "transcriptionally silent", *i.e.*, a TLCC-2 gene which is normally not expressed, or is expressed only at very low levels in a cell line or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, a transcriptionally silent, endogenous TLCC-2 gene may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such that it is operatively linked with an endogenous TLCC-2 gene, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described, *e.g.*, in Chappel, U.S. Patent No. 5,272,071; PCT publication No. WO 91/06667, published May 16, 1991.

II. Isolated TLCC-2 Proteins and Anti-TLCC-2 Antibodies

One aspect of the invention pertains to isolated TLCC-2 proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-TLCC-2 antibodies. In one embodiment, native TLCC-2 proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, TLCC-2 proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a TLCC-2 protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the TLCC-2 protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of TLCC-2 protein in which the protein is separated from cellular components of the cells from which it is

isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of TLCC-2 protein having less than about 30% (by dry weight) of non-TLCC-2 protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-TLCC-2 protein, still more preferably less than about 10% of non-TLCC-2 protein, and most preferably less than about 5% non-TLCC-2 protein. When the TLCC-2 protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of TLCC-2 protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of TLCC-2 protein having less than about 30% (by dry weight) of chemical precursors or non-TLCC-2 chemicals, more preferably less than about 20% chemical precursors or non-TLCC-2 chemicals, still more preferably less than about 10% chemical precursors or non-TLCC-2 chemicals, and most preferably less than about 5% chemical precursors or non-TLCC-2 chemicals.

As used herein, a "biologically active portion" of a TLCC-2 protein includes a fragment of a TLCC-2 protein which participates in an interaction between a TLCC-2 molecule and a non-TLCC-2 molecule. Biologically active portions of a TLCC-2 protein include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the TLCC-2 protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2, which include less amino acids than the full length TLCC-2 proteins, and exhibit at least one activity of a TLCC-2 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the TLCC-2 protein, *e.g.*, modulating membrane excitation mechanisms. A biologically active portion of a TLCC-2 protein can be a polypeptide which is, for example, 25, 30, 35, 40, 45, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 274, 500, 525, 550, 575, 580, or more amino acids in length. Biologically active

portions of a TLCC-2 protein can be used as targets for developing agents which modulate a TLCC-2 mediated activity, *e.g.*, a membrane excitation mechanism.

In one embodiment, a biologically active portion of a TLCC-2 protein comprises at least one transmembrane domain. It is to be understood that a preferred biologically active portion of a TLCC-2 protein of the present invention comprises at least one or more of the following domains: a transmembrane domain, an N-glycosylation site, a pore domain, and a proline rich domain. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native TLCC-2 protein.

In a preferred embodiment, the TLCC-2 protein has an amino acid sequence shown in SEQ ID NO:2. In other embodiments, the TLCC-2 protein is substantially identical to SEQ ID NO:2, and retains the functional activity of the protein of SEQ ID NO:2, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the TLCC-2 protein is a protein which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:2.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (*e.g.*, when aligning a second sequence to the TLCC-2 amino acid sequence of SEQ ID NO:2 having 580 amino acid residues, at least 50, preferably at least 100, more preferably at least 200, even more preferably at least 300, and even more preferably at least 400 or 500 or more amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide

as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (*Comput. Appl. Biosci.*, 4:11-17 (1988)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to TLCC-2 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 100, wordlength = 3, and a Blosum62 matrix to obtain amino acid sequences homologous to TLCC-2 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.*

25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov.

The invention also provides TLCC-2 chimeric or fusion proteins. As used herein, a TLCC-2 "chimeric protein" or "fusion protein" comprises a TLCC-2 polypeptide operatively linked to a non-TLCC-2 polypeptide. A "TLCC-2 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to TLCC-2, whereas a "non-TLCC-2 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the TLCC-2 protein, *e.g.*, a protein which is different from the TLCC-2 protein and which is derived from the same or a different organism. Within a TLCC-2 fusion protein the TLCC-2 polypeptide can correspond to all or a portion of a TLCC-2 protein. In a preferred embodiment, a TLCC-2 fusion protein comprises at least one biologically active portion of a TLCC-2 protein. In another preferred embodiment, a TLCC-2 fusion protein comprises at least two biologically active portions of a TLCC-2 protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the TLCC-2 polypeptide and the non-TLCC-2 polypeptide are fused in-frame to each other. The non-TLCC-2 polypeptide can be fused to the N-terminus or C-terminus of the TLCC-2 polypeptide.

For example, in one embodiment, the fusion protein is a GST-TLCC-2 fusion protein in which the TLCC-2 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant TLCC-2.

In another embodiment, the fusion protein is a TLCC-2 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of TLCC-2 can be increased through the use of a heterologous signal sequence.

The TLCC-2 fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The TLCC-2 fusion proteins can be used to affect the bioavailability of a TLCC-2 substrate. Use of TLCC-2 fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a TLCC-2 protein;

(ii) mis-regulation of the TLCC-2 gene; and (iii) aberrant post-translational modification of a TLCC-2 protein.

Moreover, the TLCC-2-fusion proteins of the invention can be used as immunogens to produce anti-TLCC-2 antibodies in a subject, to purify TLCC-2 ligands and in screening assays to identify molecules which inhibit the interaction of TLCC-2 with a TLCC-2 substrate.

Preferably, a TLCC-2 chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A TLCC-2-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the TLCC-2 protein.

The present invention also pertains to variants of the TLCC-2 proteins which function as either TLCC-2 agonists (mimetics) or as TLCC-2 antagonists. Variants of the TLCC-2 proteins can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a TLCC-2 protein. An agonist of the TLCC-2 proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a TLCC-2 protein. An antagonist of a TLCC-2 protein can inhibit one or more of the activities of the naturally occurring form of the TLCC-2 protein by, for example, competitively modulating a TLCC-2-mediated activity of a TLCC-2 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the

biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the TLCC-2 protein.

In one embodiment, variants of a TLCC-2 protein which function as either TLCC-2 agonists (mimetics) or as TLCC-2 antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a TLCC-2 protein for TLCC-2 protein agonist or antagonist activity. In one embodiment, a variegated library of TLCC-2 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of TLCC-2 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential TLCC-2 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of TLCC-2 sequences therein. There are a variety of methods which can be used to produce libraries of potential TLCC-2 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential TLCC-2 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477.

In addition, libraries of fragments of a TLCC-2 protein coding sequence can be used to generate a variegated population of TLCC-2 fragments for screening and subsequent selection of variants of a TLCC-2 protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a TLCC-2 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which

encodes N-terminal, C-terminal and internal fragments of various sizes of the TLCC-2 protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of TLCC-2 proteins. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify TLCC-2 variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

In one embodiment, cell based assays can be exploited to analyze a variegated TLCC-2 library. For example, a library of expression vectors can be transfected into a cell line, *e.g.*, an endothelial cell line, which ordinarily responds to TLCC-2 in a particular TLCC-2 substrate-dependent manner. The transfected cells are then contacted with TLCC-2 and the effect of expression of the mutant on signaling by the TLCC-2 substrate can be detected, *e.g.*, by monitoring intracellular calcium, IP3, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, or the activity of a TLCC-2-regulated transcription factor. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the TLCC-2 substrate, and the individual clones further characterized.

An isolated TLCC-2 protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind TLCC-2 using standard techniques for polyclonal and monoclonal antibody preparation. A full-length TLCC-2 protein can be used or, alternatively, the invention provides antigenic peptide fragments of TLCC-2 for use as immunogens. The antigenic peptide of TLCC-2 comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2 and encompasses an

epitope of TLCC-2 such that an antibody raised against the peptide forms a specific immune complex with TLCC-2. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of TLCC-2 that are located on the surface of the protein, *e.g.*, hydrophilic regions, as well as regions with high antigenicity (see, for example, Figure 2).

A TLCC-2 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed TLCC-2 protein or a chemically synthesized TLCC-2 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic TLCC-2 preparation induces a polyclonal anti-TLCC-2 antibody response.

Accordingly, another aspect of the invention pertains to anti-TLCC-2 antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as TLCC-2. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind TLCC-2. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of TLCC-2. A monoclonal antibody composition thus typically displays a single binding affinity for a particular TLCC-2 protein with which it immunoreacts.

Polyclonal anti-TLCC-2 antibodies can be prepared as described above by immunizing a suitable subject with a TLCC-2 immunogen. The anti-TLCC-2 antibody titer in the immunized subject can be monitored over time by standard techniques, such

as with an enzyme linked immunosorbent assay (ELISA) using immobilized TLCC-2. If desired, the antibody molecules directed against TLCC-2 can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the anti-TLCC-2 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem* .255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally R. H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 54:387-402; M. L. Gefter *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a TLCC-2 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds TLCC-2.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-TLCC-2 monoclonal antibody (see, *e.g.*, G. Galfre *et al.* (1977) *Nature* 266:55052; Gefter *et al.* *Somatic Cell Genet.*, cited *supra*; Lerner, *Yale J. Biol. Med.*, cited *supra*; Kenneth, *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing

hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind TLCC-2, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-TLCC-2 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with TLCC-2 to thereby isolate immunoglobulin library members that bind TLCC-2. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92/15679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al. Nature* (1990) 348:552-554.

Additionally, recombinant anti-TLCC-2 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-TLCC-2 antibody (*e.g.*, monoclonal antibody) can be used to isolate TLCC-2 by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-TLCC-2 antibody can facilitate the purification of natural TLCC-2 from cells and of recombinantly produced TLCC-2 expressed in host cells. Moreover, an anti-TLCC-2 antibody can be used to detect TLCC-2 protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the TLCC-2 protein. Anti-TLCC-2 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, ß-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable

fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a TLCC-2 protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of

interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, TLCC-2 proteins, mutant forms of TLCC-2 proteins, fusion proteins, and the like).

The recombinant expression vectors of the invention can be designed for expression of TLCC-2 proteins in prokaryotic or eukaryotic cells. For example, TLCC-2 proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion

moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be utilized in TLCC-2 activity assays, (*e.g.*, direct assays or competitive assays described in detail below), or to generate antibodies specific for TLCC-2 proteins, for example. In a preferred embodiment, a TLCC-2 fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (*e.g.*, six (6) weeks).

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118).

Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the TLCC-2 expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

Alternatively, TLCC-2 proteins can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters

(*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to TLCC-2 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a TLCC-2 nucleic acid molecule of the invention is introduced, *e.g.*, a TLCC-2 nucleic acid molecule within a recombinant expression vector or a TLCC-2 nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or

environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a TLCC-2 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a TLCC-2 protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) a TLCC-2 protein. Accordingly, the invention further provides methods for producing a TLCC-2 protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a TLCC-2 protein has been introduced) in a suitable medium such that a TLCC-2 protein is produced. In

another embodiment, the method further comprises isolating a TLCC-2 protein from the medium or the host cell.

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which TLCC-2-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous TLCC-2 sequences have been introduced into their genome or homologous recombinant animals in which endogenous TLCC-2 sequences have been altered. Such animals are useful for studying the function and/or activity of a TLCC-2 and for identifying and/or evaluating modulators of TLCC-2 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous TLCC-2 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a TLCC-2-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The TLCC-2 cDNA sequence of SEQ ID NO:1 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human TLCC-2 gene, such as a mouse or rat TLCC-2 gene, can be used as a transgene. Alternatively, a TLCC-2 gene homologue, such as another TLCC-2 family member, can be isolated based on hybridization to the TLCC-2 cDNA sequences of SEQ ID NO:1 or 3, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____ (described further in subsection I above) and used as a

transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a TLCC-2 transgene to direct expression of a TLCC-2 protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a TLCC-2 transgene in its genome and/or expression of TLCC-2 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a TLCC-2 protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a TLCC-2 gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the TLCC-2 gene. The TLCC-2 gene can be a human gene (*e.g.*, the cDNA of SEQ ID NO:3), but more preferably, is a non-human homologue of a human TLCC-2 gene (*e.g.*, a cDNA isolated by stringent hybridization with the nucleotide sequence of SEQ ID NO:1). For example, a mouse TLCC-2 gene can be used to construct a homologous recombination nucleic acid molecule, *e.g.*, a vector, suitable for altering an endogenous TLCC-2 gene in the mouse genome. In a preferred embodiment, the homologous recombination nucleic acid molecule is designed such that, upon homologous recombination, the endogenous TLCC-2 gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the homologous recombination nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous TLCC-2 gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous TLCC-2 protein). In the homologous recombination nucleic acid molecule, the altered portion of the TLCC-2 gene is flanked at its 5' and 3'

ends by additional nucleic acid sequence of the TLCC-2 gene to allow for homologous recombination to occur between the exogenous TLCC-2 gene carried by the homologous recombination nucleic acid molecule and an endogenous TLCC-2 gene in a cell, *e.g.*, an embryonic stem cell. The additional flanking TLCC-2 nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the homologous recombination nucleic acid molecule (see, *e.g.*, Thomas, K.R. and Capecchi, M. R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid molecule is introduced into a cell, *e.g.*, an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced TLCC-2 gene has homologously recombined with the endogenous TLCC-2 gene are selected (see *e.g.*, Li, E. *et al.* (1992) *Cell* 69:915). The selected cells can then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see *e.g.*, Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, *e.g.*, vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and WO 93/04169 by Berns *et al.*

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a

selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0 phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

IV. Pharmaceutical Compositions

The TLCC-2 nucleic acid molecules, fragments of TLCC-2 proteins, and anti-TLCC-2 antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions

used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a fragment of a TLCC-2 protein or an anti-TLCC-2 antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the

use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While

compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

In a preferred example, a subject is treated with antibody, protein, or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody, protein, or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in

dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (*i.e*,. including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention.

Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (*e.g.*, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram.

It isfurthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the

activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

Further, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon *et al.*, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld *et al.* (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera *et al.* (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin *et al.* (eds.), pp. 303-16 (Academic Press 1985), and Thorpe *et al.*, "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982). Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells; *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (*e.g.*, therapeutic and

prophylactic). As described herein, a TLCC-2 protein of the invention has one or more of the following activities: (1) modulates membrane excitability, (2) influences the resting potential of membranes, (3) modulates wave forms and frequencies of action potentials, (4) modulates thresholds of excitation, (5) modulates neurite outgrowth and synaptogenesis, (6) modulates signal transduction, and (7) participates in nociception.

The isolated nucleic acid molecules of the invention can be used, for example, to express TLCC-2 protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect TLCC-2 mRNA (*e.g.*, in a biological sample) or a genetic alteration in a TLCC-2 gene, and to modulate TLCC-2 activity, as described further below. The TLCC-2 proteins can be used to treat disorders characterized by insufficient or excessive production of a TLCC-2 substrate or production of TLCC-2 inhibitors. In addition, the TLCC-2 proteins can be used to screen for naturally occurring TLCC-2 substrates, to screen for drugs or compounds which modulate TLCC-2 activity, as well as to treat disorders characterized by insufficient or excessive production of TLCC-2 protein or production of TLCC-2 protein forms which have decreased, aberrant or unwanted activity compared to TLCC-2 wild type protein (*e.g.*, CNS disorders (such as neurodegenerative disorders), pain disorders, or disorders of cellular growth, differentiation, or migration. Moreover, the anti-TLCC-2 antibodies of the invention can be used to detect and isolate TLCC-2 proteins, to regulate the bioavailability of TLCC-2 proteins, and modulate TLCC-2 activity.

A. Screening Assays:

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) which bind to TLCC-2 proteins, have a stimulatory or inhibitory effect on, for example, TLCC-2 expression or TLCC-2 activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of TLCC-2 substrate.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a TLCC-2 protein or polypeptide or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a

TLCC-2 protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a TLCC-2 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate TLCC-2 activity is determined. Determining the ability of the test compound to modulate TLCC-2 activity can be accomplished by monitoring, for example, intracellular calcium, IP3, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, or the activity of a TLCC-2-regulated transcription factor. The cell, for example, can be of mammalian origin, *e.g.*, a neuronal cell, or a liver cell.

The ability of the test compound to modulate TLCC-2 binding to a substrate or to bind to TLCC-2 can also be determined. Determining the ability of the test compound to modulate TLCC-2 binding to a substrate can be accomplished, for example, by coupling the TLCC-2 substrate with a radioisotope or enzymatic label such that binding of the TLCC-2 substrate to TLCC-2 can be determined by detecting the labeled TLCC-2 substrate in a complex. Alternatively, TLCC-2 could be coupled with a radioisotope or enzymatic label to monitor the ability of a test compound to modulate TLCC-2 binding to a TLCC-2 substrate in a complex. Determining the ability of the test compound to bind TLCC-2 can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to TLCC-2 can be determined by detecting the labeled TLCC-2 compound in a complex. For example, compounds (*e.g.*, TLCC-2 substrates) can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a compound (*e.g.*, a TLCC-2 substrate) to interact with TLCC-2 without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with TLCC-2 without the labeling of either the compound or the TLCC-2. McConnell, H. M. *et al.* (1992) *Science* 257:1906-1912. As used herein, a "microphysiometer" (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and TLCC-2.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a TLCC-2 target molecule (*e.g.*, a TLCC-2 substrate) with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the TLCC-2 target molecule. Determining the ability of the test compound to modulate the activity of a TLCC-2 target molecule can be

accomplished, for example, by determining the ability of the TLCC-2 protein to bind to or interact with the TLCC-2 target molecule.

Determining the ability of the TLCC-2 protein, or a biologically active fragment thereof, to bind to or interact with a TLCC-2 target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the TLCC-2 protein to bind to or interact with a TLCC-2 target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.*, intracellular Ca^{2+}, diacylglycerol, IP_3, and the like), detecting catalytic/enzymatic activity of the target using an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a target-regulated cellular response.

In yet another embodiment, an assay of the present invention is a cell-free assay in which a TLCC-2 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the TLCC-2 protein or biologically active portion thereof is determined. Preferred biologically active portions of the TLCC-2 proteins to be used in assays of the present invention include fragments which participate in interactions with non-TLCC-2 molecules, *e.g.*, fragments with high surface probability scores (see, for example, Figure 2). Binding of the test compound to the TLCC-2 protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the TLCC-2 protein or biologically active portion thereof with a known compound which binds TLCC-2 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TLCC-2 protein, wherein determining the ability of the test compound to interact with a TLCC-2 protein comprises determining the ability of the test compound to preferentially bind to TLCC-2 or biologically active portion thereof as compared to the known compound.

In another embodiment, the assay is a cell-free assay in which a TLCC-2 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the TLCC-2 protein or biologically active portion thereof is determined. Determining the ability of

the test compound to modulate the activity of a TLCC-2 protein can be accomplished, for example, by determining the ability of the TLCC-2 protein to bind to a TLCC-2 target molecule by one of the methods described above for determining direct binding. Determining the ability of the TLCC-2 protein to bind to a TLCC-2 target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In an alternative embodiment, determining the ability of the test compound to modulate the activity of a TLCC-2 protein can be accomplished by determining the ability of the TLCC-2 protein to further modulate the activity of a downstream effector of a TLCC-2 target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

In yet another embodiment, the cell-free assay involves contacting a TLCC-2 protein or biologically active portion thereof with a known compound which binds the TLCC-2 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the TLCC-2 protein, wherein determining the ability of the test compound to interact with the TLCC-2 protein comprises determining the ability of the TLCC-2 protein to preferentially bind to or modulate the activity of a TLCC-2 target molecule.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either TLCC-2 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a TLCC-2 protein, or interaction of a TLCC-2 protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion

protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/ TLCC-2 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or TLCC-2 protein, and the mixture incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of TLCC-2 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a TLCC-2 protein or a TLCC-2 target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated TLCC-2 protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with TLCC-2 protein or target molecules but which do not interfere with binding of the TLCC-2 protein to its target molecule can be derivatized to the wells of the plate, and unbound target or TLCC-2 protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the TLCC-2 protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the TLCC-2 protein or target molecule.

In another embodiment, modulators of TLCC-2 expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of TLCC-2 mRNA or protein in the cell is determined. The level of expression of TLCC-2 mRNA or protein in the presence of the candidate compound is compared to the level of expression of TLCC-2 mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of TLCC-2 expression

based on this comparison. For example, when expression of TLCC-2 mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of TLCC-2 mRNA or protein expression. Alternatively, when expression of TLCC-2 mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of TLCC-2 mRNA or protein expression. The level of TLCC-2 mRNA or protein expression in the cells can be determined by methods described herein for detecting TLCC-2 mRNA or protein.

In yet another aspect of the invention, the TLCC-2 proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with TLCC-2 ("TLCC-2-binding proteins" or "TLCC-2-bp") and are involved in TLCC-2 activity. Such TLCC-2-binding proteins are also likely to be involved in the propagation of signals by the TLCC-2 proteins or TLCC-2 targets as, for example, downstream elements of a TLCC-2-mediated signaling pathway. Alternatively, such TLCC-2-binding proteins are likely to be TLCC-2 inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a TLCC-2 protein is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a TLCC-2-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell

colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the TLCC-2 protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a TLCC-2 protein can be confirmed *in vivo*, *e.g.*, in an animal such as an animal model for cellular pain.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a TLCC-2 modulating agent, an antisense TLCC-2 nucleic acid molecule, a TLCC-2-specific antibody, or a TLCC-2-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

Models for studying pain *in vivo* include rat models of neuropathic pain caused by methods such as intraperitoneal administration of Taxol (Authier *et al.* (2000) *Brain Res.* 887:239-249), chronic constriction injury (CCI), partial sciatic nerve transection (Linenlaub and Sommer (2000) *Pain* 89:97-106), transection of the tibial and sural nerves (Lee *et al.* (2000) *Neurosci. Lett.* 291:29-32), the spared nerve injury model (Decosterd and Woolf (2000) *Pain* 87:149-158), cuffing the sciatic nerve (Pitcher and Henry (2000) *Eur. J. Neurosci.* 12:2006-2020), unilateral tight ligation (Esser and Sawynok (2000) *Eur. J. Pharmacol.* 399:131-139), L5 spinal nerve ligation (Honroe *et al.* (2000) *Neurosci.* 98:585-598), and photochemically induced ischemic nerve injury (Hao *et al.* (2000) *Exp. Neurol.* 163:231-238); rat models of nociceptive pain caused by methods such as the Chung Method, the Bennett Method, and intraperitoneal administration of complete Freund's adjuvant (CFA) (Abdi *et al.* (2000) *Anesth. Analg.* 91:955-959); rat models of post-incisional pain caused by incising the skin and fascia of a hind paw (Olivera and Prado (2000) *Braz. J. Med. Biol. Res.* 33:957-960); rat models of cancer pain caused by methods such as injecting osteolytic sarcoma cells into the

femur (Honroe *et al.* (2000) *Neurosci.* 98:585-598); and rat models of visceral pain caused by methods such as intraperitoneal administration of cyclophosphamide.

Various methods of determining an animal's response to pain are known in the art. Examples of such methods include, but are not limited to brief intense exposure to a focused heat source, administration of a noxious chemical subcutaneously, the tail flick test, the hot plate test, the formalin test, Von Frey threshold, and testing for stress-induced analgesia (*et al.*, by restraint, foot shock, and/or cold water swim) (Crawley (2000) *What's Wrong With My Mouse?* Wiley-Liss pp. 72-75).

B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

1. Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the TLCC-2 nucleotide sequences, described herein, can be used to map the location of the TLCC-2 genes on a chromosome. The mapping of the TLCC-2 sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, TLCC-2 genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the TLCC-2 nucleotide sequences. Computer analysis of the TLCC-2 sequences can be used to predict primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the TLCC-2 sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but human cells can, the one human chromosome that contains the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. (D'Eustachio P. *et al.* (1983) *Science* 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the TLCC-2 nucleotide sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a TLCC-2 sequence to its chromosome include *in situ* hybridization (described in Fan, Y. *et al.* (1990) *Proc. Natl. Acad. Sci. USA*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical such as colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this

technique, see Verma *et al.*, Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland, J. *et al.* (1987) *Nature,* 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the TLCC-2 gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

2. Tissue Typing

The TLCC-2 sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations

of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the TLCC-2 nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The TLCC-2 nucleotide sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from TLCC-2 nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

3. Use of TLCC-2 Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, *e.g.*, hair or skin, or body fluids, *e.g.*, blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the TLCC-2 nucleotide sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of SEQ ID NO:1 having a length of at least 20 bases, preferably at least 30 bases.

The TLCC-2 nucleotide sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such TLCC-2 probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, TLCC-2 primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

C. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining TLCC-2 protein and/or nucleic acid expression as well as TLCC-2 activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant or unwanted TLCC-2 expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with TLCC-2 protein, nucleic acid expression or activity. For example, mutations in a TLCC-2 gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby phophylactically treat an individual prior to the onset of a disorder characterized by or associated with TLCC-2 protein, nucleic acid expression or activity.

Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of TLCC-2 in clinical trials.

These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

An exemplary method for detecting the presence or absence of TLCC-2 protein or nucleic acid in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting TLCC-2 protein or nucleic acid (*e.g.*, mRNA, or genomic DNA) that encodes TLCC-2 protein such that the presence of TLCC-2 protein or nucleic acid is detected in the biological sample. A preferred agent for detecting TLCC-2 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to TLCC-2 mRNA or genomic DNA. The nucleic acid probe can be, for example, the TLCC-2 nucleic acid set forth in SEQ ID NO:1 or 3, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize

under stringent conditions to TLCC-2 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting TLCC-2 protein is an antibody capable of binding to TLCC-2 protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')2) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect TLCC-2 mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of TLCC-2 mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of TLCC-2 protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of TLCC-2 genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of TLCC-2 protein include introducing into a subject a labeled anti-TLCC-2 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting TLCC-2 protein, mRNA, or genomic DNA, such that the presence of TLCC-2 protein, mRNA or genomic DNA is detected in the

biological sample, and comparing the presence of TLCC-2 protein, mRNA or genomic DNA in the control sample with the presence of TLCC-2 protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of TLCC-2 in a biological sample. For example, the kit can comprise a labeled compound or agent capable of detecting TLCC-2 protein or mRNA in a biological sample; means for determining the amount of TLCC-2 in the sample; and means for comparing the amount of TLCC-2 in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect TLCC-2 protein or nucleic acid.

2. Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant or unwanted TLCC-2 expression or activity. As used herein, the term "aberrant" includes a TLCC-2 expression or activity which deviates from the wild type TLCC-2 expression or activity. Aberrant expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant TLCC-2 expression or activity is intended to include the cases in which a mutation in the TLCC-2 gene causes the TLCC-2 gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional TLCC-2 protein or a protein which does not function in a wild-type fashion, *e.g.*, a protein which does not interact with a TLCC-2 substrate, *e.g.*, a non-calcium channel subunit or ligand and/or a non-vanilloid receptor subunit or ligand, or one which interacts with a non-TLCC-2 substrate, *e.g.* a non-calcium channel subunit or ligand and/or a non-vanilloid receptor subunit or ligand. As used herein, the term "unwanted" includes an unwanted phenomenon involved in a biological response, such as cellular proliferation. For example, the term unwanted includes a TLCC-2 expression or activity which is undesirable in a subject.

The assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with a misregulation in TLCC-2 protein activity or nucleic acid expression, such as a CNS disorder (*e.g.*, a neurodegenerative disorder, a pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder). Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disorder associated with a misregulation in TLCC-2 protein activity or nucleic acid expression, such as a CNS disorder, a pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant or unwanted TLCC-2 expression or activity in which a test sample is obtained from a subject and TLCC-2 protein or nucleic acid (*e.g.*, mRNA or genomic DNA) is detected, wherein the presence of TLCC-2 protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted TLCC-2 expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, *e.g.*, neuronal cells, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted TLCC-2 expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a CNS disorder, pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant or unwanted TLCC-2 expression or activity in which a test sample is obtained and TLCC-2 protein or nucleic acid expression or activity is detected (*e.g.*, wherein the abundance of TLCC-2 protein or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant or unwanted TLCC-2 expression or activity).

The methods of the invention can also be used to detect genetic alterations in a TLCC-2 gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in TLCC-2 protein activity or nucleic acid expression, such as a CNS disorder, pain disorder, or a disorder of cellular growth, differentiation, or migration. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a TLCC-2 -protein, or the mis-expression of the TLCC-2 gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a TLCC-2 gene; 2) an addition of one or more nucleotides to a TLCC-2 gene; 3) a substitution of one or more nucleotides of a TLCC-2 gene, 4) a chromosomal rearrangement of a TLCC-2 gene; 5) an alteration in the level of a messenger RNA transcript of a TLCC-2 gene, 6) aberrant modification of a TLCC-2 gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a TLCC-2 gene, 8) a non-wild type level of a TLCC-2-protein, 9) allelic loss of a TLCC-2 gene, and 10) inappropriate post-translational modification of a TLCC-2-protein. As described herein, there are a large number of assays known in the art which can be used for detecting alterations in a TLCC-2 gene. A preferred biological sample is a tissue or serum sample isolated by conventional means from a subject.

In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the TLCC-2-gene (see Abravaya *et al.* (1995) *Nucleic Acids Res* .23:675-682). This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a TLCC-2 gene under conditions such that hybridization and amplification of the TLCC-2-gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of

the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.* (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a TLCC-2 gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in TLCC-2 can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in TLCC-2 can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, M.T. *et al. supra.* Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is

composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the TLCC-2 gene and detect mutations by comparing the sequence of the sample TLCC-2 with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the TLCC-2 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type TLCC-2 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in TLCC-2 cDNAs obtained from samples of cells. For example, the mutY enzyme of E. coli cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a TLCC-2 sequence, *e.g.*, a wild-type TLCC-2 sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in TLCC-2 genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA*: 86:2766, see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control TLCC-2 nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is

used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl Acad. Sci USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a TLCC-2 gene.

Furthermore, any cell type or tissue in which TLCC-2 is expressed may be utilized in the prognostic assays described herein.

3. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs) on the expression or activity of a TLCC-2 protein (*e.g.*, the modulation of membrane excitability) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase TLCC-2 gene expression, protein levels, or upregulate TLCC-2 activity, can be monitored in clinical trials of subjects exhibiting decreased TLCC-2 gene expression, protein levels, or downregulated TLCC-2 activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease TLCC-2 gene expression, protein levels, or downregulate TLCC-2 activity, can be monitored in clinical trials of subjects exhibiting increased TLCC-2 gene expression, protein levels, or upregulated TLCC-2 activity. In such clinical trials, the expression or activity of a TLCC-2 gene, and preferably, other genes that have been implicated in, for example, a TLCC-2-associated disorder can be used as a "read out" or markers of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including TLCC-2, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) which modulates TLCC-2 activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on TLCC-2-associated disorders (*e.g.*, disorders characterized by deregulated signaling or membrane excitation), for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of TLCC-2 and other genes implicated in the TLCC-2-associated disorder, respectively. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of TLCC-2 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, antibody, protein, peptide, nucleic acid, ribozyme, small

molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a TLCC-2 protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the TLCC-2 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the TLCC-2 protein, mRNA, or genomic DNA in the pre-administration sample with the TLCC-2 protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of TLCC-2 to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of TLCC-2 to lower levels than detected, *i.e.* to decrease the effectiveness of the agent. According to such an embodiment, TLCC-2 expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

D. Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant or unwanted TLCC-2 expression or activity, *e.g.* a CNS disorder, pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder.

"Treatment", or "treating" as used herein, is defined as the application or administration of a therapeutic agent to a patient, or application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disease or disorder, a symptom of disease or disorder or a predisposition toward a disease or disorder, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect the disease or disorder, the symptoms of the disease or disorder, or the predisposition toward disease. A therapeutic agent includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides.

With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how a patient's genes determine his or her response to a drug (*e.g.*, a patient's "drug response phenotype", or "drug response genotype"). Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the TLCC-2 molecules of the present invention or TLCC-2 modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant or unwanted TLCC-2 expression or activity, by administering to the subject a TLCC-2 or an agent which modulates TLCC-2 expression or at least one TLCC-2 activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted TLCC-2 expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the TLCC-2 aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of TLCC-2 aberrancy, for example, a TLCC-2, TLCC-2 agonist or TLCC-2 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating TLCC-2 expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a

TLCC-2 or agent that modulates one or more of the activities of TLCC-2 protein activity associated with the cell. An agent that modulates TLCC-2 protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a TLCC-2 protein (*e.g.*, a TLCC-2 substrate), a TLCC-2 antibody, a TLCC-2 agonist or antagonist, a peptidomimetic of a TLCC-2 agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more TLCC-2 activities. Examples of such stimulatory agents include active TLCC-2 protein and a nucleic acid molecule encoding TLCC-2 that has been introduced into the cell. In another embodiment, the agent inhibits one or more TLCC-2 activities. Examples of such inhibitory agents include antisense TLCC-2 nucleic acid molecules, ribozymes, anti-TLCC-2 antibodies, and TLCC-2 inhibitors. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a TLCC-2 protein or nucleic acid molecule, *e.g.*, a pain disorder. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) TLCC-2 expression or activity. In another embodiment, the method involves administering a TLCC-2 protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted TLCC-2 expression or activity.

Stimulation of TLCC-2 activity is desirable in situations in which TLCC-2 is abnormally downregulated and/or in which increased TLCC-2 activity is likely to have a beneficial effect. Likewise, inhibition of TLCC-2 activity is desirable in situations in which TLCC-2 is abnormally upregulated and/or in which decreased TLCC-2 activity is likely to have a beneficial effect, *e.g.*, in pain disorders.

3. Pharmacogenomics

The TLCC-2 molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on TLCC-2 activity (*e.g.*, TLCC-2 gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) TLCC-2-associated disorders

(*e.g.*, proliferative disorders) associated with aberrant or unwanted TLCC-2 activity. In conjunction with such treatment, pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a TLCC-2 molecule or TLCC-2 modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a TLCC-2 molecule or TLCC-2 modulator.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11): 983-985 and Linder, M.W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (*e.g.*, a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten-million known

single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drugs target is known (*e.g.*, a TLCC-2 protein of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (*e.g.*, a TLCC-2 molecule or TLCC-2 modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a TLCC-2 molecule or TLCC-2 modulator, such as a modulator identified by one of the exemplary screening assays described herein.

VI. Electronic Apparatus Readable Media and Arrays

Electronic apparatus readable media comprising TLCC-2 sequence information is also provided. As used herein, "TLCC-2 sequence information" refers to any nucleotide and/or amino acid sequence information particular to the TLCC-2 molecules of the present invention, including but not limited to full-length nucleotide and/or amino acid sequences, partial nucleotide and/or amino acid sequences, polymorphic sequences including single nucleotide polymorphisms (SNPs), epitope sequences, and the like. Moreover, information "related to" said TLCC-2 sequence information includes detection of the presence or absence of a sequence (*e.g.*, detection of expression of a sequence, fragment, polymorphism, etc.), determination of the level of a sequence (*e.g.*, detection of a level of expression, for example, a quantative detection), detection of a reactivity to a sequence (*e.g.*, detection of protein expression and/or levels, for example, using a sequence-specific antibody), and the like. As used herein, "electronic apparatus readable media" refers to any suitable medium for storing, holding or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like; general hard disks and hybrids of these categories such as magnetic/optical storage

media. The medium is adapted or configured for having recorded thereon TLCC-2 sequence information of the present invention.

As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the TLCC-2 sequence information.

A variety of software programs and formats can be used to store the sequence information on the electronic apparatus readable medium. For example, the sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of data processor structuring formats (*e.g.*, text file or database) may be employed in order to obtain or create a medium having recorded thereon the TLCC-2 sequence information.

By providing TLCC-2 sequence information in readable form, one can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the sequence information in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has a TLCC-2- associated disease or disorder or a pre-disposition to a TLCC-2-associated disease or disorder, wherein the method comprises the steps of determining TLCC-2 sequence information associated with the subject and based on the TLCC-2 sequence information, determining

whether the subject has a TLCC-2 -associated disease or disorder or a pre-disposition to a TLCC-2-associated disease or disorder and/or recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has a TLCC-2-associated disease or disorder or a pre-disposition to a disease associated with a TLCC-2 wherein the method comprises the steps of determining TLCC-2 sequence information associated with the subject, and based on the TLCC-2 sequence information, determining whether the subject has a TLCC-2 -associated disease or disorder or a pre-disposition to a TLCC-2 -associated disease or disorder, and/or recommending a particular treatment for the disease, disorder or pre-disease condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

The present invention also provides in a network, a method for determining whether a subject has a TLCC-2 -associated disease or disorder or a pre-disposition to a TLCC-2 -associated disease or disorder associated with TLCC-2, said method comprising the steps of receiving TLCC-2 sequence information from the subject and/or information related thereto, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to TLCC-2 and/or a [TLCC-2}-associated disease or disorder, and based on one or more of the phenotypic information, the TLCC-2 information (*e.g.*, sequence information and/or information related thereto), and the acquired information, determining whether the subject has a TLCC-2-associated disease or disorder or a pre-disposition to a TLCC-2-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention also provides a business method for determining whether a subject has a TLCC-2-associated disease or disorder or a pre-disposition to a TLCC-2-associated disease or disorder, said method comprising the steps of receiving information related to TLCC-2 (*e.g.*, sequence information and/or information related thereto), receiving phenotypic information associated with the subject, acquiring information from the network related to TLCC-2 and/or related to a TLCC-2-associated disease or disorder, and based on one or more of the phenotypic information, the TLCC-

2 information, and the acquired information, determining whether the subject has a TLCC-2-associated disease or disorder or a pre-disposition to a TLCC-2-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The invention also includes an array comprising a TLCC-2 sequence of the present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression, one of which can be TLCC-2. This allows a profile to be developed showing a battery of genes specifically expressed in one or more tissues.

In addition to such qualitative determination, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression *per se* and level of expression in that tissue. This is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of a TLCC-2-associated disease or disorder, progression of TLCC-2-associated disease or disorder, and processes, such a cellular transformation associated with the TLCC-2-associated disease or disorder.

The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells (*e.g.*, ascertaining the effect of TLCC-2 expression on the expression of other genes). This provides, for example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes (*e.g.*, including TLCC-2) that could serve as a molecular target for diagnosis or therapeutic intervention.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application, as well as the Figures and the Sequence Listing, are incorporated herein by reference.

EXAMPLES

EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF HUMAN TLCC-2 cDNA

In this example, the identification and characterization of the gene encoding human TLCC-2 (clone Fbh54420FL) is described.

Isolation of the TLCC-2 cDNA

The invention is based, at least in part, on the discovery of a human gene encoding a novel protein, referred to herein as TLCC-2. The entire sequence of the human clone Fbh54420FL was determined and found to contain an open reading frame termed human "TLCC-2". The nucleotide sequence of the human TLCC-2 gene is set forth in Figures 1 and in the Sequence Listing as SEQ ID NO:1 and 3. The amino acid sequence of the human TLCC-2 expression product is set forth in Figures 1 and in the Sequence Listing as SEQ ID NO: 2.

The nucleotide sequence encoding the human TLCC-2 protein is shown in Figure 1 and is set forth as SEQ ID NO:1. The protein encoded by this nucleic acid comprises about 580 amino acids and has the amino acid sequence shown in Figure 1 and set forth as SEQ ID NO:2. The coding region (open reading frame) of SEQ ID NO:1 is set forth as SEQ ID NO:3. Clone Fbh54420FL, comprising the coding region of human TLCC-2, was deposited with the American Type Culture Collection (ATCC®), 10801 University Boulevard, Manassas, VA 20110-2209, on _____, and assigned Accession No. _____.

Analysis of the Human TLCC-2 Molecules

A BLASTN 2.0 search against the dbEST database, using a score of 100 and a wordlength of 12 (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human TLCC-2 revealed that human IC54420 is 99% identical to tb24a12.x1 NCI_CGAP_kid12 *Homo sapiens* cDNA clone IMAGE:2055262 3' similar to WP:R13A5.1 CE01370 (Accession Number AI307240) over nucleotides 2077 to 1377. The search further revealed that human TLCC-2 is 98% identical to au44h03.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone IMAGE:2517653 3' similar to WP:R13A5.1 CE01370 (Accession Number AI816064) over nucleotides 2079-1375. This search further revealed that human TLCC-2 is 97% identical to wv36f01.x1 NCI_CGAP_Ov18 *Homo sapiens* cDNA clone IMAGE:2531641 3' similar to WP:R13A5.1 CE01370 (Accession Number AI951554) over nucleotides 2088 to 1407. This search further revealed that human TLCC-2 is 97% identical to wp80f10.x1 NCI_CGAP_Brn25 *Homo sapiens* cDNA clone IMAGE:2468107 3' similar to WP:R13A5.1 CE01370 (Accession Number AI942492) over nucleotides 2072-1422. The search further revealed that human TLCC-2 is 96% identical to nr72c11.s1 NCI_CGAP_Pr24 Homo sapiens cDNA clone IMAGE:1173524 similar to WP:R13A5.1 CE01370 (Accession Number AA641031) over nucleotides 2073-1407.

A BLASTX 2.0 search against the NRP/protot database, using a score of 100 and a wordlength of 3 (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403), of the translated nucleotide sequence of human TLCC-2 revealed that human TLCC-2 is 67% or less identical to fragments (*e.g.*, fragments of 185 amino acids or less) of unnamed protein product from *Homo sapiens* (Accession Number AK001868). The search further

revealed that human TLCC-2 is 27% or less identical to fragments (*e.g.*, fragments of 75 amino acids or less) of human polycystic kidney disease and receptor for egg jelly related protein (Accession Number AF116458). This search further revealed that human TLCC-2 is 43% or less identical to fragments (*e.g.*, fragments of 59 amino acids or less) of polycystic kidney disease and receptor for egg jelly related protein from *Mus musculus* (Accession Number AF116459). This search further revealed that human TLCC-2 is 70% or less identical to fragments (*e.g.*, fragments of 29 amino acids or less) of a 110 amino acid long hypothetical protein from *Aeropyrum pernix* (Accession Number AP000060) over the full length of this protein.

An alignment of the human TLCC-2 amino acid sequence with the amino acid sequences of vanilloid receptor 1 and vanilloid receptor 2 from *Homo sapiens* (Accession Numbers AAD26363 and AAD41724, SEQ ID NO:4 and SEQ ID NO:5, respectively) using the CLUSTAL W (1.74) multiple sequence alignment program is set forth in Figure 6. An alignment of the human TLCC-2 amino acid sequence with the amino acid sequence of polycystic kidney disease protein 2 from *Mus musculus* (Accession Number NP 032887, SEQ ID NO:6) using the CLUSTAL W(1.74) multiple sequence alignment program is set forth in Figure 7A-B. An alignment of the human TLCC-2 amino acid sequence with the amino acid sequence of human melastatin (Accession Number AAC8000, SEQ ID NO:7) using the CLUSTAL W(1.74) multiple sequence alignment program is set forth in Figure 8A-B.

A search was performed against the Memsat database (Figures 2 and 3), resulting in the identification of six transmembrane domains in the amino acid sequence of human TLCC-2 (SEQ ID NO:2) at about residues 70-86, 299-317, 354-371, 385-416, 428-447, and 497-521.

A search was also performed against the Prosite database (Figure 9) resulting in the identification of four N-glycosylation sites in the amino acid sequence of human TLCC-2 (SEQ ID NO:2) at about residues 159-162, 179-182, 220-223, and 230-233.

A search was also performed against the HMM database (Figure 4) resulting in the identification of a fibronectin type III domain in the amino acid sequence of human TLCC-2 (SEQ ID NO:2) at about residues 202-269 (score = 5).

A search was also performed against the ProDom database resulting in the identification of a 31K RNA-4 protein domain in the amino acid sequence of human TLCC-2 (SEQ ID NO:2) at about residues 397-443 (score = 73). The results of the search are set forth in Figure 5.

Tissue Distribution of Human TLCC-2 mRNA by Northern Analysis

This example describes the tissue distribution of TLCC-2 mRNA, as determined by Northern analysis.

Northern blot hybridizations with the various RNA samples are performed under standard conditions and washed under stringent conditions, *i.e.*, 0.2xSSC at 65°C. The DNA probe is radioactively labeled with ^{32}P-dCTP using the Prime-It kit (Stratagene, La Jolla, CA) according to the instructions of the supplier. Filters containing human mRNA (MultiTissue Northern I and MultiTissue Northern II from Clontech, Palo Alto, CA) are probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

Tissue Distribution of TLCC-2 mRNA by *In situ* Analysis

For *in situ* analysis, various tissues, *e.g.* tissues obtained from brain and spinal cord from monkey and rat, were first frozen on dry ice. Ten-micrometer-thick sections of the tissues were post-fixed with 4% formaldehyde in DEPC treated 1X phosphate-buffered saline at room temperature for 10 minutes before being rinsed twice in DEPC 1X phosphate-buffered saline and once in 0.1 M triethanolamine-HCl (pH 8.0). Following incubation in 0.25% acetic anhydride-0.1 M triethanolamine-HCl for 10 minutes, sections were rinsed in DEPC 2X SSC (1X SSC is 0.15M NaCl plus 0.015M sodium citrate). Tissue was then dehydrated through a series of ethanol washes, incubated in 100% chloroform for 5 minutes, and then rinsed in 100% ethanol for 1 minute and 95% ethanol for 1 minute and allowed to air dry.

Hybridizations were performed with ^{35}S-radiolabeled (5 X 10^7 cpm/ml) cRNA probes. Probes were incubated in the presence of a solution containing 600 mM NaCl, 10 mM Tris (pH 7.5), 1 mM EDTA, 0.01% sheared salmon sperm DNA, 0.01% yeast tRNA, 0.05% yeast total RNA type X1, 1X Denhardt's solution, 50% formamide, 10%

dextran sulfate, 100 mM dithiothreitol, 0.1% sodium dodecyl sulfate (SDS), and 0.1% sodium thiosulfate for 18 hours at 55°C.

After hybridization, slides were washed with 2X SSC. Sections are then sequentially incubated at 37°C in TNE (a solution containing 10 mM Tris-HCl (pH 7.6), 500 mM NaCl, and 1 mM EDTA), for 10 minutes, in TNE with 10μg of RNase A per ml for 30 minutes, and finally in TNE for 10 minutes. Slides were then rinsed with 2X SSC at room temperature, washed with 2X SSC at 50°C for 1 hour, washed with 0.2X SSC at 55°C for 1 hour, and 0.2X SSC at 60°C for 1 hour. Sections were then dehydrated rapidly through serial ethanol-0.3 M sodium acetate concentrations before being air dried and exposed to Kodak Biomax MR scientific imaging film for 24 hours and subsequently dipped in NB-2 photoemulsion and exposed at 4°C for 7 days before being developed and counter stained.

In situ hybridization results showed expression in monkey and rat brain (including cortex, stratium, and hippocampus), spinal cord, and dorsal root ganglia (DRG) neurons.

Is situ hybridization in rat animal models showed up-regulation of the TLCC-2 gene 10 days after unilateral chronic constriction injury (CCI). There was up-regulation of TLCC-2 seven days after axotomy and after intraplantar injection of complete Freund's adjuvant (CFA). These levels decreased to normal levels at later time points. No contralateral effects were observed. These results indicate that the TLCC-2 molecules of the present invention are up-regulated in response to painful simuli, and are therefore involved in nociception. Modulation, *e.g.*, inhibition, of expression or activity of the TLCC-2 molecules of the invention may therefore modulate nociception and provide treatment for pain disorders.

Tissue Expression Analysis of TLCC-2 mRNA Using Taqman Analysis

This example describes the tissue distribution of human TLCC-2 mRNA in a variety of cells and tissues, as determined using the TaqMan™ procedure. The Taqman™ procedure is a quantitative, reverse transcription PCR-based approach for detecting mRNA. The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold™ DNA Polymerase to cleave a TaqMan™ probe during PCR. Briefly, cDNA was generated from the samples of interest, *e.g.*, human brain, spinal cord, heart, kidney,

liver, lung, dorsal root ganglia, and skin, and used as the starting material for PCR amplification. In addition to the 5' and 3' gene-specific primers, a gene-specific oligonucleotide probe (complementary to the region being amplified) was included in the reaction (*i.e.*, the Taqman™ probe). The TaqMan™ probe includes the oligonucleotide with a fluorescent reporter dye covalently linked to the 5' end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N',N'-tetramethylrhodamine) at the 3' end of the probe.

During the PCR reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites. The 5'-3' nucleolytic activity of the AmpliTaq™ Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target. The probe fragments are then displaced from the target, and polymerization of the strand continues. The 3' end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product. RNA was prepared using the trizol method and treated with DNase to remove contaminating genomic DNA. cDNA was synthesized using standard techniques. Mock cDNA synthesis in the absence of reverse transcriptase resulted in samples with no detectable PCR amplification of the control gene confirms efficient removal of genomic DNA contamination.

Two human normal tissue panels indicated broad distribution of human TLCC-2 expression, with highest expression in human brain, followed by testis, placenta, adrenal gland, spinal cord, skin, and dorsal root ganglia (DRG) (See Figures 10 and 11).

EXAMPLE 2: EXPRESSION OF RECOMBINANT IC54420 PROTEIN IN BACTERIAL CELLS

In this example, TLCC-2 is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, TLCC-2 is fused to GST and this fusion polypeptide is expressed in *E. coli*, *e.g.*, strain PEB199. Expression of the GST-TLCC-2 fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

EXAMPLE 3: EXPRESSION OF RECOMBINANT IC54420 PROTEIN IN COS CELLS

To express the TLCC-2 gene in COS cells, the pcDNA/Amp vector by Invitrogen Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire TLCC-2 protein and an HA tag (Wilson *et al.* (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

To construct the plasmid, the TLCC-2 DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the TLCC-2 coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the TLCC-2 coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly,

MA). Preferably the two restriction sites chosen are different so that the TLCC-2 gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5□, SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the TLCC-2-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the IC54420 polypeptide is detected by radiolabelling (^{35}S-methionine or ^{35}S-cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labelled for 8 hours with ^{35}S-methionine (or ^{35}S-cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the TLCC-2 coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the TLCC-2 polypeptide is detected by radiolabelling and immunoprecipitation using a TLCC-2-specific monoclonal antibody.

EXAMPLE 4: REGULATION OF CALCIUM INFLUX THROUGH TLCC-2

This experiment describes the regulation of calcium influx though TLCC-2 in HEK293 cells as determined by Fluorometric Imaging Plate Reader experiments (FLIPR) (Molecular Devices Corp., Sunnyvale, CA).

The FLIPR is a screening tool for cell-based fluorescent assays which allows the simultaneous stimulation and measurement of separate cell populations in a high throughput format. Therefore, using this system, it is possible to quantify transient signals, such as the release of intracellular calcium, from cell populations, in parallel and in real time. The FLIPR contains chambers in which to hold the test plate and plates containing antagonists or agonists to be added to the test plate. The FLIPR utilizes an argon laser that provides discrete spectral lines spaced from approximately 350 to 530 nm. For use with fluorescent Ca^{2+} dyes, the 88-nm line of the laser is employed. The laser simultaneously illuminates the wells in a test plate. The image of each well in the plate is captured by a cooled charge coupled device (CCD) camera, which updates images once per second, if required, for the measurement of rapid calcium responses. Because both excitation and emission are read via the bottom of the plate, black-walled, transparent bottomed 96-well plates are used. Data captured by the CCD camera is converted to digital data and then transferred to a computer.

Briefly, a calcium indicator (*e.g.*, fluo-3/AM or Calcium Green-1/AM) was transferred to the culture medium. Because the FLIPR collects fluorescence from the bottom of the well, suspension cells require centrifugation to the base of the well following dye loading. Viable HEK293 cells were resuspended in loading medium and incubated for one hour. The cells were then centrifuged and resuspended with wash buffer. The cell suspension containing the dye was then aliquotted into each well of the black-walled, transparent bottomed 96-well plate and the plate was centrifuged. The FLIPR assay was then carried out and the results analyzed. (If adherent cells are used, they may be plated at an appropriate density in the 96-well plates and cultured overnight. Dye may then be loaded and incubated).

Results show a constitutive calcium influx through TLCC-2 in HEK293 cells that were incubated with NMDG/0 Ca^{+2} and stimulated afterwards with 5mM Ca^{+2}.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

What is claimed:

1. An isolated nucleic acid molecule selected from the group consisting of:

(a) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:1;

(b) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:3;

(c) a nucleic acid molecule comprising the nucleotide sequence contained in the plasmid deposited with ATCC® as Accession Number _____;

(d) a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3;

(e) a nucleic acid molecule comprising a fragment of at least 706 nucleotides of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3;

(f) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2;

(g) a nucleic acid molecule which encodes a polypeptide comprising an amino acid sequence at least about 60% identical to the amino acid sequence of SEQ ID NO:2;

(h) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 18 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2;

(i) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2;

(j) a nucleic acid molecule comprising a nucleotide sequence which is complementary to the nucleotide sequence of the nucleic acid molecule of any one of subparts (a) to (i); and

(k) a nucleic acid molecule which hybridizes to a complement of the nucleic acid molecule of any one of subparts (a) to (d), (f), (g) and (i) under stringent conditions.

2. An isolated nucleic acid molecule comprising the nucleic acid molecule of claim 1, and a nucleotide sequence encoding a heterologous polypeptide.

3. A vector comprising the nucleic acid molecule of claim 1.

4. The vector of claim 3, which is an expression vector.

5. A host cell transfected with the expression vector of claim 4.

6. A method of producing a polypeptide comprising culturing the host cell of claim 5 in an appropriate culture medium to, thereby, produce the polypeptide.

7. An isolated polypeptide selected from the group consisting of:

(a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 18 contiguous amino acids of SEQ ID NO:2;

(b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of SEQ ID NO:1 or SEQ ID NO:3 under stringent conditions;

(c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3;

(d) a polypeptide comprising an amino acid sequence which is at least 60% identical to the amino acid sequence of SEQ ID NO:2; and

(e) a polypeptide comprising the amino acid sequence of SEQ ID NO:2.

8. The polypeptide of claim 7, further comprising heterologous amino acid sequences.

9. An antibody which selectively binds to a polypeptide of claim 7.

10. A method for detecting the presence of a polypeptide of claim 7 in a sample comprising:

(a) contacting the sample with a compound which selectively binds to the polypeptide; and

(b) determining whether the compound binds to the polypeptide in the sample to thereby detect the presence of a polypeptide of claim 7 in the sample.

11. The method of claim 10, wherein the compound which binds to the polypeptide is an antibody.

12. A kit comprising a compound which selectively binds to a polypeptide of claim 7 and instructions for use.

13. A method for detecting the presence of a nucleic acid molecule of claim 1 in a sample comprising:

(a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and

(b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample to thereby detect the presence of a nucleic acid molecule of claim 1 in the sample.

14. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of claim 1 and instructions for use.

15. A method for identifying a compound which binds to or modulates the activity of a polypeptide of claim 7 comprising:

(a) contacting the polypeptide, or a cell expressing the polypeptide with a test compound; and

(b) determining whether the compound binds to or modulates the activity of the polypeptide, to thereby identify a compound which binds to or modulates the activity of the polypeptide.

16. A method of identifying a nucleic acid molecule associated with a pain disorder comprising:

(a) contacting a sample comprising nucleic acid molecules with a hybridization probe comprising at least 25 contiguous nucleotides of SEQ ID NO:1; and

(b) detecting the presence of a nucleic acid molecule in said sample that hybridizes to said probe, thereby identifying a nucleic acid molecule associated with a pain disorder.

17. A method of identifying a nucleic acid associated with a pain disorder comprising:

(a) contacting a sample comprising nucleic acid molecules with a first and a second amplification primer, said first primer comprising at least 25 contiguous nucleotides of SEQ ID NO:1 and said second primer comprising at least 25 contiguous nucleotides from the complement of SEQ ID NO:1;

(b) incubating said sample under conditions that allow nucleic acid amplification; and

(c) detecting the presence of a nucleic acid molecule in said sample that is amplified, thereby identifying a nucleic acid molecule associated with a pain disorder.

18. A method of identifying a polypeptide associated with a pain disorder comprising:

(a) contacting a sample comprising polypeptides with a TLCC-2 binding substance; and

(b) detecting the presence of a polypeptide in said sample that binds to said TLCC-2 binding substance, thereby identifying a polypeptide associated with a pain disorder.

19. The method of claim 18, wherein said binding substance is an antibody.

20. The method of claim 18, wherein said binding substance is detectably labeled.

21. A method for identifying a compound capable of treating a pain disorder comprising assaying the ability of the compound or agent to modulate TLCC-2 expression or activity, thereby identifying a compound capable of treating a pain disorder.

22. The method of claim 21, wherein said compound inhibits TLCC-2 expression or activity.

23. A method for identifying a subject suffering from a pain disorder comprising obtaining a biological sample from the subject, and detecting in the sample aberrant or abnormal TLCC-2 expression or activity, thereby identifying a subject suffering from a pain disorder.

24. A method for treating a subject having a pain disorder characterized by aberrant TLCC-2 polypeptide activity or aberrant TLCC-2 nucleic acid expression comprising administering to the subject a TLCC-2 modulator, thereby treating said subject having a pain disorder.

25. The method of claim 24, wherein the TLCC-2 modulator is selected from the group consisting of a small molecule, an antibody specific for TLCC-2, a TLCC-2 polypeptide, a fragment of a TLCC-2 polypeptide, a TLCC-2 nucleic acid molecule, a fragment of a TLCC-2 nucleic acid molecule, an antisense TLCC-2 nucleic acid molecule, and a ribozyme.

26. The method of claim 24, wherein said TLCC-2 modulator is administered in a pharmaceutically acceptable formulation.

27. The method of claim 24, wherein said TLCC-2 modulator is administered using a gene therapy vector.

28. A method for identifying a compound capable of treating a pain disorder characterized by aberrant TLCC-2 nucleic acid expression or TLCC-2 polypeptide activity comprising assaying the ability of the compound to modulate TLCC-2 nucleic acid expression or TLCC-2 polypeptide activity, thereby identifying a compound capable of treating a pain disorder characterized by aberrant TLCC-2 nucleic acid expression or TLCC-2 polypeptide activity.

29. A method for identifying a compound capable of modulating nociception comprising:

(a) contacting a cell with a test compound; and

(b) assaying the ability of the test compound to modulate the expression of a TLCC-2 nucleic acid or the activity of a TLCC-2 polypeptide;

thereby identifying a compound capable of modulating nociception.

30. The method of claim 29, wherein the TLCC-2 modulator is a TLCC-2 inhibitor.

Input file Fbh54420FL.seq; Output File 54420.trans
Sequence length 2095

```
GTCGACCCACGCGTCCGCGGACGCGTGGGCGGAGGGTTTGAAGCCGCGCCGCGAGGGAGCGAGGTCGCAGTGACAGCGG

                                                                  M   T   A   P      4
CGGGCGATCGGACCCAGGCTGCCCCGCCGTACCCGCCTGCGTCCCGCGCTCCCGCCCCAGC ATG ACA GCC CCG    12

 A   G   P   R   G   S   E   T   E   R   L   L   T   P   N   P   G   Y   G   T     24
GCG GGT CCG CGC GGC TCA GAG ACC GAG CGG CTT CTG ACC CCC AAC CCC GGG TAT GGG ACC    72

 Q   A   G   P   S   P   A   P   P   T   P   P   E   E   E   D   L   R   R   R     44
CAG GCG GGG CCT TCA CCG GCC CCT CCG ACA CCC CCA GAA GAG GAA GAC CTT CGC CGT CGT   132

 L   K   Y   F   F   M   S   P   C   D   K   F   R   A   K   G   R   K   P   C     64
CTC AAA TAC TTT TTC ATG AGT CCC TGC GAC AAG TTT CGA GCC AAG GGC CGC AAG CCC TGC   192

 K   L   M   L   Q   V   V   K   I   L   V   V   T   V   Q   L   I   L   F   G     84
AAG CTG ATG CTG CAA GTG GTC AAG ATC CTG GTG GTC ACG GTG CAG CTC ATC CTG TTT GGG   252

 L   S   N   Q   L   A   V   T   F   R   E   E   N   T   I   A   F   R   H   L    104
CTC AGT AAT CAG CTG GCT GTG ACA TTC CGG GAA GAG AAC ACC ATC GCC TTC CGA CAC CTC   312

 F   L   L   G   Y   S   D   G   A   D   D   T   F   A   A   Y   T   R   E   Q    124
TTC CTG CTG GGC TAC TCG GAC GGA GCG GAT GAC ACC TTC GCA GCC TAC ACG CGG GAG CAG   372

 L   Y   Q   A   I   F   H   A   V   D   Q   Y   L   A   L   P   D   V   S   L    144
CTG TAC CAG GCC ATC TTC CAT GCT GTG GAC CAG TAC CTG GCG TTG CCT GAC GTG TCA CTG   432

 G   R   Y   A   Y   V   R   G   G   G   D   P   W   T   N   G   S   G   L   A    164
GGC CGG TAT GCG TAT GTC CGT GGT GGG GGT GAC CCT TGG ACC AAT GGC TCA GGG CTT GCT   492

 L   C   Q   R   Y   Y   H   R   G   H   V   D   P   A   N   D   T   F   D   I    184
CTC TGC CAG CGG TAC TAC CAC CGA GGC CAC GTG GAC CCG GCC AAC GAC ACA TTT GAC ATT   552

 D   P   M   V   V   T   D   C   I   Q   V   D   P   P   E   R   P   P   P   P    204
GAT CCG ATG GTG GTT ACT GAC TGC ATC CAG GTG GAT CCC CCC GAG CGG CCC CCT CCG CCC   612

 P   S   D   D   L   T   L   L   E   S   S   S   S   Y   K   N   L   T   L   K    224
CCC AGC GAC GAT CTC ACC CTC TTG GAA AGC AGC TCC AGT TAC AAG AAC CTC ACG CTC AAA   672

 F   H   K   L   V   N   V   T   I   H   F   R   L   K   T   I   N   L   Q   S    244
TTC CAC AAG CTG GTC AAT GTC ACC ATC CAC TTC CGG CTG AAG ACC ATT AAC CTC CAG AGC   732

 L   I   N   N   E   I   P   D   C   Y   T   F   S   V   L   I   T   F   D   N    264
CTC ATC AAT AAT GAG ATC CCG GAC TGC TAT ACC TTC AGC GTC CTG ATC ACG TTT GAC AAC   792

 K   A   H   S   G   R   I   P   I   S   L   E   T   Q   A   H   I   Q   E   C    284
AAA GCA CAC AGT GGG CGG ATC CCC ATC AGC CTG GAG ACC CAG GCC CAC ATC CAG GAG TGT   852

 K   H   P   S   V   F   Q   H   G   D   N   S   F   R   L   L   F   D   V   V    304
AAG CAC CCC AGT GTC TTC CAG CAC GGA GAC AAC AGC TTC CGG CTC CTG TTT GAC GTG GTG   912

 V   I   L   T   C   S   L   S   F   L   L   C   A   R   S   L   L   R   G   F    324
GTC ATC CTC ACC TGC TCC CTG TCC TTC CTC CTC TGC GCC CGC TCA CTC CTT CGA GGC TTC   972

 L   L   Q   N   E   F   V   G   F   M   W   R   Q   R   G   R   V   I   S   L    344
CTG CTG CAG AAC GAG TTT GTG GGG TTC ATG TGG CGG CAG CGG GGA CGG GTC ATC AGC CTG  1032

 W   E   R   L   E   F   V   N   G   W   Y   I   L   L   V   T   S   D   V   L    364
TGG GAG CGG CTG GAA TTT GTC AAT GGC TGG TAC ATC CTG CTC GTC ACC AGC GAT GTG CTC  1092
```

FIGURE 1A

```
 T  I  S  G  T  I  M  K  I  G  I  E  A  K  N  L  A  S  Y  D   384
ACC ATC TCG GGC ACC ATC ATG AAG ATC GGC ATC GAG GCC AAG AAC TTG GCG AGC TAC GAC  1152

 V  C  S  I  L  L  G  T  S  T  L  L  V  W  V  G  V  I  R  Y   404
GTC TGC AGC ATC CTC CTG GGC ACC TCG ACG CTG CTG GTG TGG GTG GGC GTG ATC CGC TAC  1212

 L  T  F  F  H  N  Y  N  I  L  I  A  T  L  R  V  A  L  P  S   424
CTG ACC TTC TTC CAC AAC TAC AAT ATC CTC ATC GCC ACA CTG CGG GTG GCC CTG CCC AGC  1272

 V  M  R  F  C  C  C  V  A  V  I  Y  L  G  Y  C  F  C  G  W   444
GTC ATG CGC TTC TGC TGC TGC GTG GCT GTC ATC TAC CTG GGC TAC TGC TTC TGT GGC TGG  1332

 I  V  L  G  P  Y  H  V  K  F  R  S  L  S  M  V  S  E  C  L   464
ATC GTG CTG GGG CCC TAT CAT GTG AAG TTC CGC TCA CTC TCC ATG GTG TCT GAG TGC CTG  1392

 F  S  L  I  N  G  D  D  M  F  V  T  F  A  A  M  Q  A  Q  Q   484
TTC TCG CTC ATC AAT GGG GAC GAC ATG TTT GTG ACG TTC GCC GCC ATG CAG GCG CAG CAG  1452

 G  R  S  S  L  V  W  L  F  S  Q  L  Y  L  Y  S  F  I  S  L   504
GGC CGC AGC AGC CTG GTG TGG CTC TTC TCC CAG CTC TAC CTT TAC TCC TTC ATC AGC CTC  1512

 F  I  Y  M  V  L  S  L  F  I  A  L  I  T  G  A  Y  D  T  I   524
TTC ATC TAC ATG GTG CTC AGC CTC TTC ATC GCG CTC ATC ACC GGC GCC TAC GAC ACC ATC  1572

 K  H  P  G  G  A  G  A  E  E  S  E  L  Q  A  Y  I  A  Q  C   544
AAG CAT CCC GGC GGC GCA GGC GCA GAG GAG AGC GAG CTG CAG GCC TAC ATC GCA CAG TGC  1632

 Q  D  S  P  T  S  G  K  F  R  R  G  S  G  S  A  C  S  L  L   564
CAG GAC AGC CCC ACC TCC GGC AAG TTC CGC CGC GGG AGC GGC TCG GCC TGC AGC CTT CTC  1692

 C  C  C  G  R  D  P  S  E  E  H  S  L  L  V  N  *            581
TGC TGC TGC GGA AGG GAC CCC TCG GAG GAG CAT TCG CTG CTG GTG AAT TGA              1743

TTCGACCTGACTGCCGTTGGACCGTAGGCCCTGGACTGCAGAGACCCCCGCCCCGACCCCGCTTATTTATTTGTAGGGT

TTGCTTTTAAGGATCGGCTCCCTGTCGCGCCCGAGGAGGGCCTGGACCTTTCGTGTCGGACCCTTGGGGGCGGGGAGAC

TGGGTGGGAAGGTGTGAATAAAAGGGAAATTAAAAAAAAAAAAAAAAATTTTAAA
```

FIGURE 1B

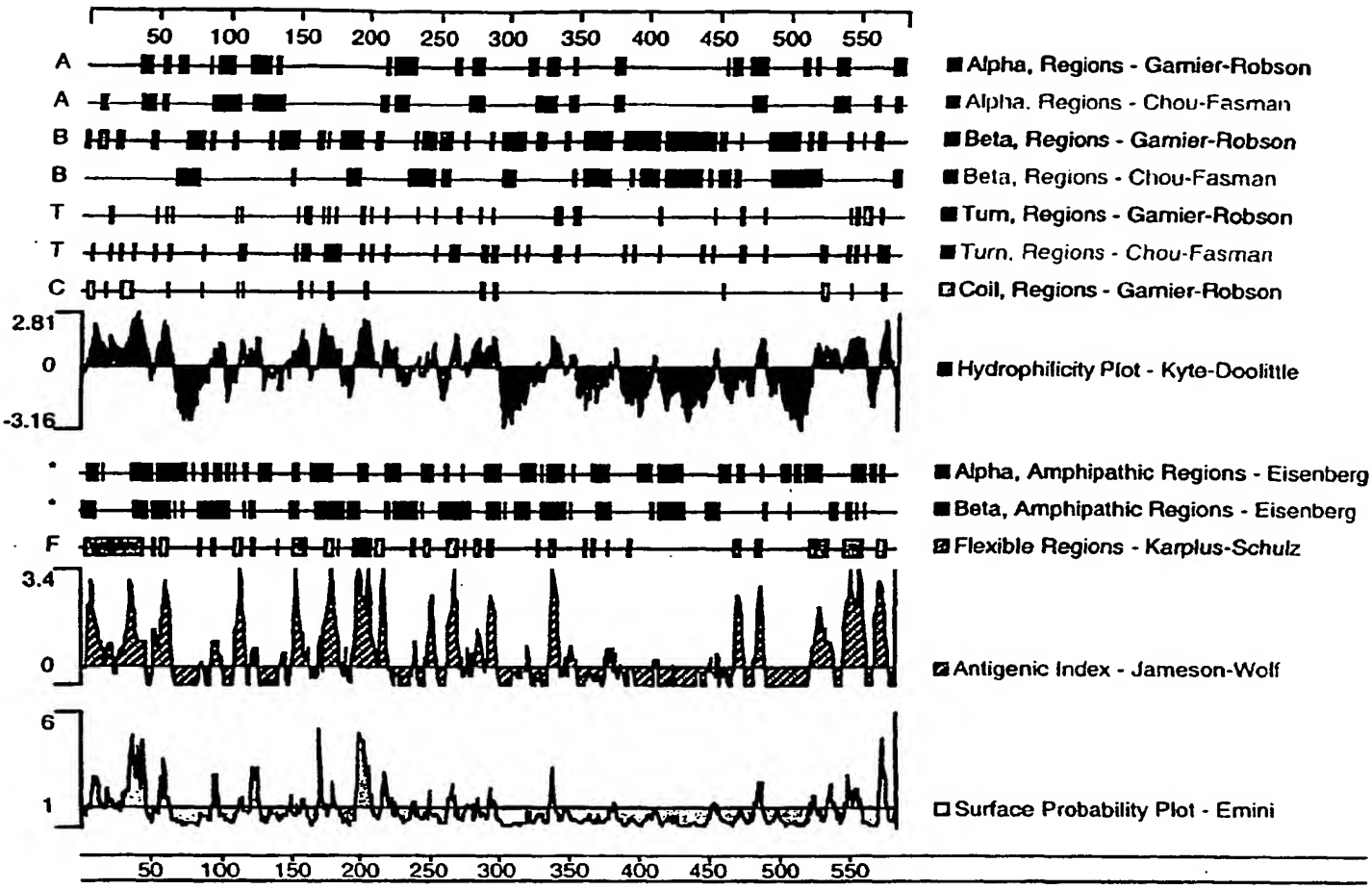


FIGURE 2

Transmembrane Segments Predicted by MEMSAT

Start	End	Orient	Score
70	86	ins-->out	2.0
299	317	out-->ins	3.7
354	371	ins-->out	2.4
385	402	out-->ins	3.1
428	447	ins-->out	4.7
497	521	out-->ins	6.4

>54420
MTAPAGPRGSETERLLTPNPGYGTQAGPSPAPPTPPEEEDLRRRLKYFFHSPCDKFRAKG
RKPCKLMLQVVKILVVTVQLILFGLSNQLAVTFREENTIAFRHLFLLGYSDGADDTFAAY
TREQLYQAIFHAVDQYLALPDVSLGRYAYVRGGGDPWTNGSGLALCQRYYHRGHVDPAND
TFDIDPHVVTDCIQVDPPERPPPPPSDDLTLLESSSSYKNLTLKFHKLVNVTIHPRLKTI
NLQSLINNEIPDCYTFSVLITFDNKAHSGRIPISLETQAHIQECKHPSVPQHGDNSFRLL
FDVVVILTCSLSFLLCARSLLRGFLLQNEFVGFMWRQRGRVISLWERLEPVNGMYILLVT
SDVLTISGTIMKIGIEAKNLASYDVCSILLGTSTLLVWVGVIRYLTPFHNYNILIATLRV
ALPSVMRFCCCVAVIYLGYCFCGWIVLGPYHVKFRSLSMVSECLFSLINGDDMFVTFAAM
QAQQGRSSLVWLFSQLYLVSFISLFIYMVLSLFIALITGAYDTIKHPGGAGAEESELQAY
IAQCQDSPTSGKFRRGSGSACSLLCCCGRDPSEEHSLLVN

FIGURE 3

Protein Family / Domain Matches, HMMer version 2

```
Searching for complete domains in PFAM
hmmpfam - search a single seq against HMM database
HMMER 2.1.1 (Dec 1998)
Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL).
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HMM file:                 /prod/ddm/seqanal/PFAM/pfam4.4/Pfam
Sequence file:            /tmp/orfanal.12622.aa
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
  Query:  54420

Scores for sequence family classification (score includes all domains):
Model    Description                                           Score    E-value  N
-------- -----------                                           -----    ------- ---
        [no hits above thresholds]

Parsed for domains:
Model    Domain  seq-f seq-t    hmm-f hmm-t       score  E-value
-------- ------- ----- -----    ----- -----       -----  -------
        [no hits above thresholds]

Alignments of top-scoring domains:
        [no hits above thresholds]
//
Searching for complete domains in SMART
hmmpfam - search a single seq against HMM database
HMMER 2.1.1 (Dec 1998)
Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL).
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HMM file:                 /ddm/robison/smart/smart/smart.all.hmms
Sequence file:            /tmp/orfanal.12622.aa
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
  Query:  54420

Scores for sequence family classification (score includes all domains):
Model    Description                                           Score    E-value  N
-------- -----------                                           -----    ------- ---
FN3_2    .                                                       5.0        9.4   1

Parsed for domains:
Model    Domain  seq-f seq-t    hmm-f hmm-t       score  E-value
-------- ------- ----- -----    ----- -----       -----  -------
FN3_2      1/1     202   269 ..     1    69 []     5.0        9.4

Alignments of top-scoring domains:
FN3_2: domain 1 of 1, from 202 to 269: score 5.0, E = 9.4
                  *->PgpPsnptelrvtdvtststsdltsvtlsWepP...ivgYr......
                     P+pPs++  l++ + +s++       +tl+  +  + ++++r ++ +
       54420   202    PPPPSDD--LTLLESSSSYK----NLTLKFHXLvnvTIHFRlktinl 242

                  .ytltgLkPgaepwteYefrVrAvngndaGeg<-*
                  +  + +  P+       Y+f+V    +n+a +g
       54420   243 qSLINNEIPD-----CYTFSVLITFDNKAHSG    269
```

FIGURE 4

ProDom Matches

ProdomId	Start	End	Description	Score
View Prodom 35860	397	443	p99.2 (2) Q65670(1) Y32K(1) // PROTEIN 31K RNA-4	73
ProdomId	**Start**	**End**	**Description**	**Score**

View Prodom 35860

```
>35860 p99.2 (2) Q65670(1) Y32K(1)  // PROTEIN 31K RNA-4
       Length = 152

 Score = 73 (30.8 bits), Expect = 4.0, P = 0.98
 Identities = 15/50 (30%), Positives = 27/50 (54%)

Query:   397 VWVGVIRYLTFFHNYNILIATLRVALPSVHRFC-CCVAVIYLG--YCFCG 443
             VW+ ++   T +  +++++   +  LPS +  C-CC  V Y     +CFCG
Sbjct:    21 VWILLLTSSTCYGYHDVVVDIEQCTLPSNIDGCVCCSGVCYFNDNHCFCG 70
```

FIGURE 5

CLUSTAL W (1.74) multiple sequence alignment 7/14

```
humanVR1        MKKWSSTDLGTAADPLQKDTCPDPLDGDPNSRPPPAKPQLPTAKSRTRLFGKGDSEEAFP
hVR2            MTSPSSS-------PVFR---LETLDGGQEDGSEADRGKLD--------FGSG----LPP
54420.pro       MTAPAGP----------R--------GSETERLLTPNPGYG--------TQAG----PSP
                *.  :..          :        *.  .   . .              *      *

humanVR1        VDCPHEEGELDSCPTITVSPVITIQRPGDGPTGARLLSQDSVAASTEKTLRLYDRRSIFE
hVR2            MESQFQGEDRKFAPQIRVN---LNYRKGTG-------------ASQPDPNR-FDRDRLFN
54420.pro       APPTPPEEE-DLRRRLKYF----------------------FMSPCDKFR--------A
                         :  .    :                        *  .  *

humanVR1        AVAQNNCQDLESLLLFLQKSKKHLTDNEFKDPETGKTCLLKAMLNLHDGQNTTIPLLLEI
hVR2            AVSRGVPEDLAGLPEYLSKTSKYLTDSEYTEGSTGKTCLMKAVLNLKDGVNACILPLLQI
54420.pro       K-GRKPCKLMLQVVKILVVT---VQLILP-----GLSNQLAVTFREE---NTIAFRHLFL
                  .:   : :  :   *  :   :    :     * :  : . :. .  *:     * :

humanVR1        ARQTDSLKELVNASYTDSYYKGQTALHIAIERRNMALVTLLVENGADVQAAAHGDFFKKT
hVR2            DRDSGNPQPLVNAQCTDDYYRGHSALHIAIEKRSLQCVKLLVENGANVHARACGRFFQKG
54420.pro       LGYSDGADDTFAAYTREQLY--QAIFHAVDQYLALPDVSL----GRYAYVRGGGDPWTNG
                   :.. .  . *   :. *  :: :* . :   :  *.*.   *  . . . *  : :

humanVR1        KGRPGFYFGELPLSLAACTNQLGIVKFLLQNSWQTADISARDSVGNTVLHALVEVADNTA
hVR2            QG-TCFYFGELPLSLAACTKQWDVVSYLLENPHQPASLQATDSQGNTVLHALVMISDNSA
54420.pro       SG--------L----ALC--QR----YYHRGHVDPAN-DTFDIDP-MVVTDCIQV-DPPE
                .*        *    * *  *     :  ..  :.*. .: *     *:   : : * .

humanVR1        DNTKFVTSMYNEILMLGAKLHPTLKLEELTNKKGMTPLALAAGTGKIGVLAYILQREIQE
hVR2            ENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKLAAKEGKIEIFRHILQREFSG
54420.pro       RPPPPPSDDLTLLESSSSYKNLTLKFHKLVNVTIHFRLKTIN----------LQSLINN
                      :.    :   .:    *:::..: *      *                 **  :.

humanVR1        PECRHLSRKFTEWAYGPVHSSLYDLSCIDTCEKNSVLEVIAYSSSETPNRHDMLLVEPLN
hVR2            --LSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEIIAFH-CKSPHRHRMVVLEPLN
54420.pro       --E--IPD-----CY-----TFSVLITFDNKAHSGRIPISLET-----QAHIQECKHPS-
                     :.      .*     ::  *  .*.  ... : :          : *     .*

humanVR1        RLLQDKWDRFVKRIFYFNFLVYCLYMIIFTMAAYYRP--VDGLPPFKMEKIGDYFRVTGE
hVR2            KLLQAKWDLLIPK-FFLNFLCNLIYMFIFTAVAYHQPTLKKQAAPHLKAEVGNSMLLTGH
54420.pro       -VFQH-GDNSFRLLFDVVVILTCSLSFLLCARSLLRG-----FL-LQNEFVGFMWRQRGR
                 ::*   *  .   * . .:      :::   :  :  .           :*      *.

humanVR1        ILSVLGGVYFFFRGIQYFLQRRPSMKTLFVDSYSEMLFFLQSLFMLATVVLYFSHLKEYV
hVR2            ILILLGGIYLLVGQLWYFWRRHVFIWISFIDSYFEILFLFQALLTVVSQVLCFLAIEWYL
54420.pro       VISLWERLEFVNG--WY---------ILLVTS--DVLTISGTIMKIGIEAKNLASYDVCS
                :: :   : :.     *           :: *  ::* :  ::: :   .  :   .

humanVR1        ASMVFSLALGWTNMLYYTRGFQQMGIYAVMIEKMILRDLCRFMFVYIVFLFGFSTAVVTL
hVR2            PLLVSALVLGWLNLLYYTRGFQHTGIYSVMIQKVILRDLLRFLLIYLVFLFGFAVALVSL
54420.pro       ILLGTSTLLVWVGVIRYLTFFHNYNILIATLR-VALPSVMRFCCCVAVIYLG-------Y
                  :  : * * .:: *   *:: .*  . :. : * .: **      *: :*

humanVR1        IEDGKNDSLPSESTSHRWRGPACRPPD----SSYNSLYSTCLELFKFTIGMGDLEFTENY
hVR2            SQEAWRPEAPTGPNATESVQPMEGQEDEGNGAQYRGILEASLELFKFTIGMGELAFQEQL
54420.pro       CFCGWIVLGPY----HVKFRSLSMVSE-----CLFSLINGDDMFVTFAA-MQAQQGRSSL
                   .    *          .    :        .: .    :..*:  *       ..

humanVR1        DFKAVFIILLLAYVILTYILLLNMLIALMGETVNKIAQESKNIWKLQRAITILDTEKSFL
hVR2            HFRGMVLLLLLAYVLLTYILLLNMLIALMSETVNSVATDSWSIWKLQKAISVLEMENGYW


54420.pro       -VWLFSQLYLYSFISLFIYMVLSLFIALITGAYDTIKHPG-GAGAEESELQAYIAQ----
                 .  .  : * ::: *   ::*.::***:  : :.:   . .    :  :     :

humanVR1        KCMRKAFRSGKLLQVGYTPDGKDDYRWCFRVDEVNWTTWNTNVGIINEDPGNCEGVKRTL
hVR2            WCRKKQ-RAGVMLTVGTKPDGSPDERWCFRVEEVNWASWEQTLPTLCEDP-SGAGVPRTL
54420.pro       -CQDSP--TSGKFRRG-S--GS--A--CSLLCCCGRDPSEEHSLLVN-------------
                 *  .   :.  :  * .  *.     *  :   .  . :      :

humanVR1        SFSLRSSRVSGRHWKNFALVPLLREASARDRQSAQPEEVYLRQFSGSLKPEDAEVFKSPA
hVR2            EN-------------------PVLASPPKEDEDGASEENYVPVQLLQSN------------
54420.pro       ------------------------------------------------------------

humanVR1        ASGEK
hVR2            -----
54420.pro       -----
```

FIGURE 6

```
CLUSTAL W (1.74) multiple sequence alignment

54420.pro       ------------------------------------------------------------
mousePKD2       MVNSRRVQPQPPGDAGRSPAPRASGPGRLVAGGAGLAVPGGLGEQRGLEIEMERIRQAAA

54420.pro       ------------------------------------------------------MTAPAG
mousePKD2       RDPPAGASASPSPPLSSCSRQAWSRDNPGFEAEEDDDDDEVEGEEGGMVVEMDVEWRPGS
                                                                      *..

54420.pro       PR---GSETERLLTPNPGYGTQAGPSPAPPTPPEEEDLRRRLKYFFMS--PCDK-FRAKG
mousePKD2       RRSASSSAVSSVGARGRGLGSYRGAAHLSGRRRRLEDQGAQCPSPAGGGDPLHRHLPLEG
                 *   .* .. : : . * *:  *.:  .    . **   :       . * .: :  :*

54420.pro       RKP-CKLMLQVVKIL--VVTVQLILFGLSNQ---LAVTFREENTIAFRHLFLLGYSDGAD
mousePKD2       QPPRVAWAERLVRGLRGLWGTRLMEESNANREKYLKSVLRELVTYLFFLVVLCILTYGMM
                : *       ::*: *  :  .:*:  . :*:   *  .:**  *  *  :.*   : *

54420.pro       DT-FAAYTR--EQLYQAIFHAVDQYLALPDVS--LGRYAYVRG---GGDPWTNGSG--LA
mousePKD2       SSNVYYYTRTLSQLFIDTPVSKTEKTNFKTLSSMEDFWKFTEGSFLDGLYWKAQTSNHTQ
                .: .  *** .**:     :  :   :  :*   . : :..*   .*  *.  :.

54420.pro       LCQR---YYHRGHV-DP-------ANDTFDIDPMV---VTDCIQVD---PPERPPPPPSD
mousePKD2       ADNRSFIFYENLLLGVPRLRQLRVRNGSCSIPQDLRDEIKECYDVYSVSSEDRAPFGPRN
                  :*   :*..  :  *        *.: .*   :   :.:* :*    . :*.*  * :

54420.pro       DLTLLESSSSYKN--------LTLKFHKLVNVTIHFRLKTIN-LQSLINNEIPDCYTFSV
mousePKD2       GTAWMYTSEKELNGSSHWGIIASYSGAGYYLDLSRTREETAAQLAGLRRNFWLDRGTRAA
                . : : :*..  *          : .          : * :*   * .* .*   *  * :.

54420.pro       LITF--DNKAHSGRIPISLETQAHIQECKHPS-VFQHG-----DNSFRLLFDVVVILTCS
mousePKD2       FIDFSVYNANINLFCVVRLLAEFPATGGVVPSWQFQPVKLIRYVTAFDFFLAACEIIFCF
                :* *   *   .    : * ::        **  **         .:* ::: .  *: *

54420.pro       LSFLLCARSLLRGFLLQNEFVGFMWRQRGRVISLWERLEFVNGWYILLVTSDVLTISGTI
mousePKD2       FIIYYVVEEILEIRIHRLSYFRSFWNCLDVVIVVLSVVAMVINIYRMSNAEGLLQFLEDQ
                : :   ...:*.  : : .:.  :*.  . ** : . : :* . * :  :..:* :

54420.pro       MKIGIEAKNLASYDVCS-ILLGTSTLLVWVGVIRYLTFFHNYNILIATLRVALPSVMRFC
mousePKD2       NSFPN-FEHVAYWQIQFNNISAVMVFLVWIKLFKFINFNRTMSQLSTTMSRCAKDLFGFT
                 .:    :::* :::    : .. .:***: :::::.* :. . * :*:  .  .:: *

54420.pro       CCVAVIYLGYCFCGWIVLGPYHVKFRSLSMVSECLFSLINGDDMFVTFAAMQAQQGRSSL
mousePKD2       IMFSIIFLAYAQLAYLVFG---TQVDDFSTFQECIFTQFRIILGDINFAEIEEAN-R---
                  .::*:*.*.  .::*:*   .:. .:* ..**:*: :.     :.** ::  : *

54420.pro       VWLFSQLYLYSFISLFIYMVLSLFIALITGAYDTIKHPGGAGAEESELQAYIAQ-CQDSP
mousePKD2       --VLGPLYFTTFVFFMFFILLNMFLAIINDSYSEVKSDLAQQKAEMELSDLIRKGCQKAL
                  ::. **: :*: ::::::*.:*:*:*..:*. :*   .    * **.  * : **.:

54420.pro       TSGKFRRGSGSACSLLCCCG----------RDPSEE-HSLLVN----------------
mousePKD2       VKLKLKRNTVDAISESLRQGGGKLNFDELRQDLKGEGHTDAEIEAIFTKYDQDGDQELTE
                .. *::*.: .* *     *          :* . * *:

54420.pro       ------------------------------------------------------------
mousePKD2       REHQQMRDDLEKEREDLDLEHSSLPRPMSSRSFPRSLDDSEEEDDEDSGHSSRRRGSISS

54420.pro       ------------------------------------------------------------
```

FIGURE 7A

```
mousePKD2       SYEEFQVLVRRVDRMEHSIGSIVSKIDAV[redacted]KLEIMERAKLKRREVLGRLLDGVAEDA

54420.pro       ------------------------------------------------------------
mousePKD2       RLGRDSEIHREQMERLVREELERWESDDAASQTGHGVSTQVGLGGQPHPRNPRPPSSQSA

54420.pro       ----------------
mousePKD2       EGLEGGGGNGSANVHA
```

FIGURE 7B

```
CLUSTAL W (1.74) Multiple sequence alignment

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     MYIRVSYDTKPDSLLHLMVKDWQLELPKLLISVHGGLQNFEMQPKLKQVFGKGLIKAAMT

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     TGAWIFTGGVSTGVISHVGDALKDHSSKSRGRVCAIGIAPWGIVENKEDLVGKDVTRVYQ

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     TMSNPLSKLSVLNNSHTHFILADNGTLGKYGAEVKLRRLLEKHISLQKINTRLGQGVPLV

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     GLVVEGGPNVVSIVLEYLQEEPPIPVVICDGSGRASDILSFAHKYCEEGGIINESLREQL

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     LVTIQKTFNYNKAQSHQLFAIIMECMKKKELVTVFRMGSEGQQDIEMAILTALLKGTNVS

Fbh54420FL      -----MTAPAGPRGSET-ERLLTPNP---GYGTQAGPSPAPPT-----PPEE--EDLRRR
hMELASTATIN     APDQLSLALAWNRVDIARSQIFVFGPHWTPLGSLAPPTDSKATEKEKKPPMATTKGGRGK
                      * *  * . : .:::. .*     *: * *: : .*     **     :. * :

Fbh54420FL      LKYFFMSPCDKFRAKGRKPCKLMLQVVKILVVTVQLILFGLSNQLAVTFREENTIAFRHL
hMELASTATIN     GKGKKKGKVKEEVEEETDPRKIELLNWVNALEQAMLDALVLDRVDFVKLLIENGVNMQHF
                 *    .  .:   :  .* *: *      :  . *  : *..   *.:  ** : ::*:

Fbh54420FL      FLLG-----YSD--GADDTFAAYTR---EQLYQAIFHA--VDQYLALPDVSLGRY--AYV
hMELASTATIN     LTIPRLEELYNTRLGPPNTLHLLVRDVKKSNLPPDYHISLIDIGLVLEYLMGGAYRCNYT
                : :       *.   *. :*:   .*   :.   . :*   :*  *.*  :  * *   *.

Fbh54420FL      RGGG-DPWTNGSGLALCQRYYHRGHVD--PANDTFDIDPMVVTDCIQV---DPPERPPPP
hMELASTATIN     RKNFRTLYNNLFGPKRPKALKLLGMEDDEPPAKGKKKKKKKKEEEIDIDVDDPAVSRFQY
                * .    :.*  *    :     *  *  *. .  .  .     : *::   **.

Fbh54420FL      PSDDLTLLESSSSYKNLTLKFHKLVNVTIHFRLKTINLQSLINNEIPDCYTFS-VLITFD
hMELASTATIN     PFHELMVWAVLMKRQKMAVFLWQRGEESMAKALVACKLYKAMAHESSESDLVDDISQDLD
                * .:* :     . ::::: : :  : ::   * : :* . : :* .:.  .. :   :*

Fbh54420FL      N-KAHSGRIPISLETQ--AHIQECKHPSV---FQHGDNSFRLLFDVV------VILTCSL
hMELASTATIN     NNSKDFGQLALELLDQSYKHDEQIAMKLLTYELKNWSNSTCLKLAVAAKHRDFIAHTCSQ
                * . . *::.:.*  *   * ::     :   ::: .**  * : *.       :  ***

Fbh54420FL      SFL-------LCAR-----SLLRGFLLQNEFVGFMWRQRGRVISLWERLEFVNGWYILLV
hMELASTATIN     MLLTDMWMGRLRMRKNPGLKVIMGILLP-PTILFLEFRTYDDFSYQTSKENEDGKEKEEE
                 :*       * *     .:: *:**    : *:  :  .  :*     *  :*

Fbh54420FL      TSDVLTISGTIMKIG-IEAKNLASY----DVCSILLGTSTLLVWVGVIRYLTFFHNYNIL
hMELASTATIN     NTDANADAGSRKGDEENEHKKQRSIPIGTKICEFYN-APIVKFWFYTISYLGYLLLFNYV
                .:*. : :*:        * *:  *     .:*.:   :. : .*. .* ** ::  :* :

Fbh54420FL      IATLRVALPSVMRFCCCVAVIYLGYCFCGWIVLGPYHVKFRSLSMVSECLFSLINGDDM-
hMELASTATIN     ILVRMDGWPSLQEWIVISYIVSLALEKIREILMSEPGKLSQKIKVWLQEYWNITDLVAIS
                * .   . **: .:     :: *.      *::.      :.:.:  :  :.: :   :

Fbh54420FL      -FVTFAAMQAQ-Q-----GRSS-----LVWLFSQLYLYSFISLF-------------IY
```

FIGURE 8A

```
hMELASTATIN     TFMIGAILRLQNQPYMGYGRVIYCVDIIF[redacted]IRVLDIFGVNKYLGPYVMMIGKMMIDMLY
                 *:  * :: * *     **          *.:    .*  *     :   . :    ..

Fbh54420FL      MVLSLFIALITG--AYDTIKHP----GGAGAEESELQAY--IAQCQDSPTSGKFRRGSGS
hMELASTATIN     FVVIMLVVLMSFGVARQAILHPEEKPSWKLARNIFYMPYWMIYGEVFADQIDLYAMEINP
                :*: ::::.*::   * ::* **    .   *.:    .*  *     :   . :    ..

Fbh54420FL      AC--SLLCCCG-RDPSEEHSLLVN----------------------------------
hMELASTATIN     PCGENLYDEEGKRLPPCIPGAWLTPALMACYLLVANILLVNLLIAVFNNTFFEVKSISNQ
                .*  .*    * * *.   .  :.

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     VWKFQRYQLIMTFHDRPVLPPPMIILSHIYIIIMRLSGRCRKKREGDQEERDRGLKLFLS

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     DEELKRLHEFEEQCVQEHFREKEDEQQSSSDERIRVTSERVENMSMRLEEINERETFMKT

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     SLQTVDLRLAQLEELSNRMVNALENLAGIDRSDLIQARSRASSECEATYLLRQSSINSAD

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     GYSLYRYHFNGEELLFEDTSLSTSPGTGVRKKTCSFRIKEEKDVKTHLVPECQNSLHLSL

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     GTSTSATPDGSHLAVDDLKNAEESKLGPDIGISKEDDERQTDSKKEETISPSLNKTDVIH

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     GQDKSDVQNTQLTVETTNIEGTISYPLEETKITRYFPDETINACKTMKSRSFVYSRGRKL

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     VGGVNQDVEYSSITDQQLTTEWQCQVQKITRSHSTDIPYIVSEAAVQAEQKEQFADMQDE

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     HHVAEAIPRIPRLSLTITDRNGMENLLSVKPDQTLGFPSLRSKSLHGHPRNVKSIQGKLD

Fbh54420FL      ----------------------------------
hMELASTATIN     RSGHASSVSSLVIVSGMTAEEKKVKKEKASTETEC
```

FIGURE 8B

Prosite version: Release 12.2 of February 1995

>PS00001|PDOC00001|ASN_GLYCOSYLATION N-glycosylation site.

Query: 159 NGSG 162

Query: 179 NDTF 182

Query: 220 NLTL 223

Query: 230 NVTI 233

>PS00004|PDOC00004|CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation site.

Query: 554 RRGS 557

>PS00005|PDOC00005|PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 12 TER 14

Query: 92 TFR 94

Query: 217 SYK 219

Query: 222 TLK 224

Query: 268 SGR 270

Query: 296 SFR 298

Query: 417 TLR 419

Query: 523 TIK 525

Query: 550 SGK 552

>PS00006|PDOC00006|CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query: 10 SETE 13

Query: 34 TPPE 37

Query: 51 SPCD 54

Query: 92 TFRE 95

Query: 210 TLLE 213

Query: 343 SLWE 346

>PS00008|PDOC00008|MYRISTYL N-myristoylation site.

Query: 112 GADDTF 117

Query: 160 GSGLAL 165

Query: 519 GAYDTI 524

Query: 528 GGAGAE 533

Query: 556 GSGSAC 561

>PS00009|PDOC00009|AMIDATION Amidation site.

Query: 59 KGRK 62

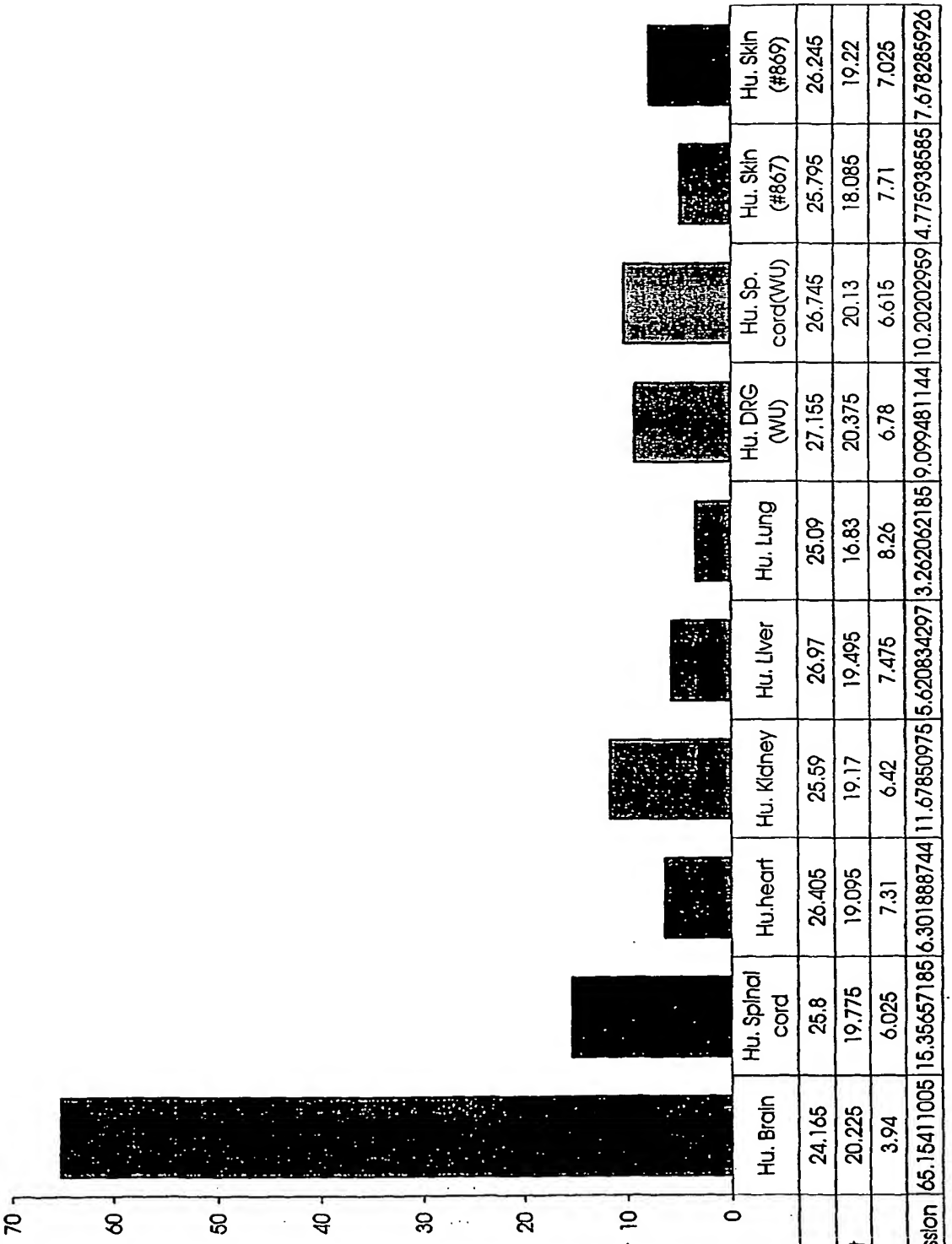
>PS00029|PDOC00029|LEUCINE_ZIPPER Leucine zipper pattern.

Query: 300 *LPDVVVILTCSLSFLLCARSLL* 321

>PS00120|PDOC00110|LIPASE_SER Lipases, serine active site.

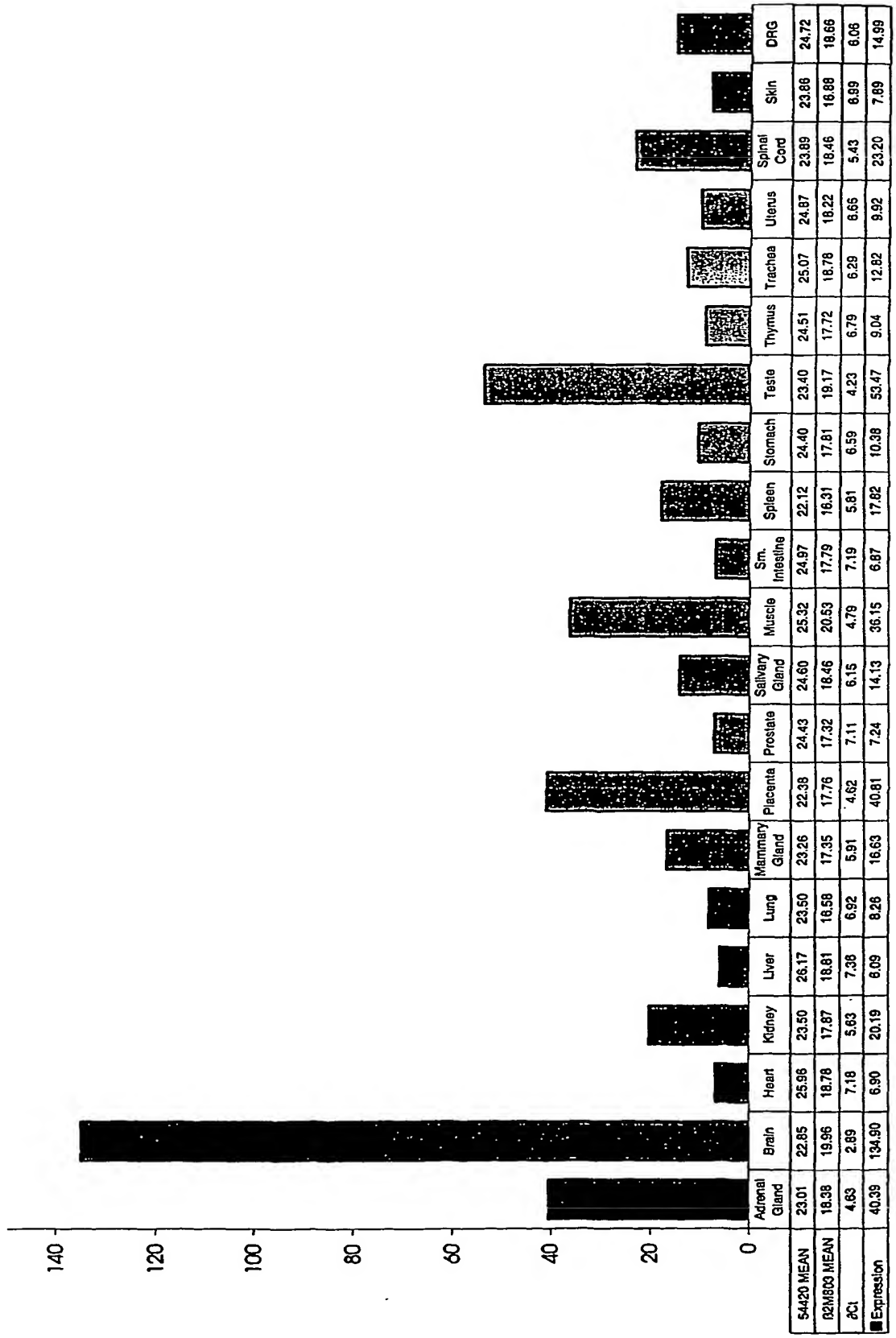
Query: 104 LFLLGYSDGA 113

FIGURE 9



	Hu. Brain	Hu. Spinal cord	Hu.heart	Hu. Kidney	Hu. Liver	Hu. Lung	Hu. DRG (WU)	Hu. Sp. cord(WU)	Hu. Skin (#867)	Hu. Skin (#869)
average Ct	24.165	25.8	26.405	25.59	26.97	25.09	27.155	26.745	25.795	26.245
average HK Ct	20.225	19.775	19.095	19.17	19.495	16.83	20.375	20.13	18.085	19.22
delta Ct	3.94	6.025	7.31	6.42	7.475	8.26	6.78	6.615	7.71	7.025
Relative Expression	65.15411005	15.35657185	6.301888744	11.67850975	5.620834297	3.262062185	9.099481144	10.20202959	4.775938585	7.678285926

FIGURE 10



	Adrenal Gland	Brain	Heart	Kidney	Liver	Lung	Mammary Gland	Placenta	Prostate	Salivary Gland	Muscle	Sm. Intestine	Spleen	Stomach	Testis	Thymus	Trachea	Uterus	Spinal Cord	Skin	DRG
54420 MEAN	23.01	22.85	25.96	23.50	26.17	23.50	23.26	22.38	24.43	24.60	25.32	24.97	22.12	24.40	23.40	24.51	25.07	24.87	23.89	23.86	24.72
02M803 MEAN	18.38	19.96	18.78	17.87	18.81	18.58	17.35	17.76	17.32	18.46	20.53	17.79	16.31	17.81	19.17	17.72	18.78	18.22	18.46	18.88	18.66
ΔCt	4.63	2.89	7.18	5.63	7.38	6.92	5.91	4.62	7.11	6.15	4.79	7.19	5.81	6.59	4.23	6.79	6.29	6.66	5.43	6.89	6.06
Expression	40.39	134.90	6.90	20.19	6.09	8.26	16.63	40.81	7.24	14.13	36.15	6.87	17.82	10.38	53.47	9.04	12.82	9.92	23.20	7.89	14.99

FIGURE 11

SEQUENCE LISTING

<110> Millennium Pharmaceuticals, Inc.

<120> 54420, A NOVEL HUMAN CALCIUM CHANNEL

<130> MNI-125CPPC

<140>
<141>

<150> US 09/544,797
<151> 2000-04-07

<160> 7

<170> PatentIn Ver. 2.0

<210> 1
<211> 2095
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (141)..(1880)

<400> 1
gtcgacccac gcgtccgcgg acgcgtgggc ggagggtttg aagccgcgcc gcgagggagc 60

gaggtcgcag tgacagcggc gggcgatcgg acccaggctg ccccgccgta cccgcctgcg 120

tcccgcgctc ccgccccagc atg aca gcc ccg gcg ggt ccg cgc ggc tca gag 173
Met Thr Ala Pro Ala Gly Pro Arg Gly Ser Glu
1 5 10

acc gag cgg ctt ctg acc ccc aac ccc ggg tat ggg acc cag gcg ggg 221
Thr Glu Arg Leu Leu Thr Pro Asn Pro Gly Tyr Gly Thr Gln Ala Gly
15 20 25

cct tca ccg gcc cct ccg aca ccc cca gaa gag gaa gac ctt cgc cgt 269
Pro Ser Pro Ala Pro Pro Thr Pro Pro Glu Glu Glu Asp Leu Arg Arg
30 35 40

cgt ctc aaa tac ttt ttc atg agt ccc tgc gac aag ttt cga gcc aag 317
Arg Leu Lys Tyr Phe Phe Met Ser Pro Cys Asp Lys Phe Arg Ala Lys
45 50 55

ggc cgc aag ccc tgc aag ctg atg ctg caa gtg gtc aag atc ctg gtg 365
Gly Arg Lys Pro Cys Lys Leu Met Leu Gln Val Val Lys Ile Leu Val
60 65 70 75

gtc acg gtg cag ctc atc ctg ttt ggg ctc agt aat cag ctg gct gtg 413
Val Thr Val Gln Leu Ile Leu Phe Gly Leu Ser Asn Gln Leu Ala Val
80 85 90

aca ttc cgg gaa gag aac acc atc gcc ttc cga cac ctc ttc ctg ctg 461
Thr Phe Arg Glu Glu Asn Thr Ile Ala Phe Arg His Leu Phe Leu Leu
95 100 105

ggc tac tcg gac gga gcg gat gac acc ttc gca gcc tac acg cgg gag 509

```
Gly Tyr Ser Asp Gly Ala Asp Asp Thr Phe Ala Ala Tyr Thr Arg Glu
        110                 115                 120

cag ctg tac cag gcc atc ttc cat gct gtg gac cag tac ctg gcg ttg   557
Gln Leu Tyr Gln Ala Ile Phe His Ala Val Asp Gln Tyr Leu Ala Leu
    125                 130                 135

cct gac gtg tca ctg ggc cgg tat gcg tat gtc cgt ggt ggg ggt gac   605
Pro Asp Val Ser Leu Gly Arg Tyr Ala Tyr Val Arg Gly Gly Gly Asp
140                 145                 150                 155

cct tgg acc aat ggc tca ggg ctt gct ctc tgc cag cgg tac tac cac   653
Pro Trp Thr Asn Gly Ser Gly Leu Ala Leu Cys Gln Arg Tyr Tyr His
                160                 165                 170

cga ggc cac gtg gac ccg gcc aac gac aca ttt gac att gat ccg atg   701
Arg Gly His Val Asp Pro Ala Asn Asp Thr Phe Asp Ile Asp Pro Met
            175                 180                 185

gtg gtt act gac tgc atc cag gtg gat ccc ccc gag cgg ccc cct ccg   749
Val Val Thr Asp Cys Ile Gln Val Asp Pro Pro Glu Arg Pro Pro Pro
        190                 195                 200

ccc ccc agc gac gat ctc acc ctc ttg gaa agc agc tcc agt tac aag   797
Pro Pro Ser Asp Asp Leu Thr Leu Leu Glu Ser Ser Ser Ser Tyr Lys
    205                 210                 215

aac ctc acg ctc aaa ttc cac aag ctg gtc aat gtc acc atc cac ttc   845
Asn Leu Thr Leu Lys Phe His Lys Leu Val Asn Val Thr Ile His Phe
220                 225                 230                 235

cgg ctg aag acc att aac ctc cag agc ctc atc aat aat gag atc ccg   893
Arg Leu Lys Thr Ile Asn Leu Gln Ser Leu Ile Asn Asn Glu Ile Pro
                240                 245                 250

gac tgc tat acc ttc agc gtc ctg atc acg ttt gac aac aaa gca cac   941
Asp Cys Tyr Thr Phe Ser Val Leu Ile Thr Phe Asp Asn Lys Ala His
            255                 260                 265

agt ggg cgg atc ccc atc agc ctg gag acc cag gcc cac atc cag gag   989
Ser Gly Arg Ile Pro Ile Ser Leu Glu Thr Gln Ala His Ile Gln Glu
        270                 275                 280

tgt aag cac ccc agt gtc ttc cag cac gga gac aac agc ttc cgg ctc   1037
Cys Lys His Pro Ser Val Phe Gln His Gly Asp Asn Ser Phe Arg Leu
    285                 290                 295

ctg ttt gac gtg gtg gtc atc ctc acc tgc tcc ctg tcc ttc ctc ctc   1085
Leu Phe Asp Val Val Val Ile Leu Thr Cys Ser Leu Ser Phe Leu Leu
300                 305                 310                 315

tgc gcc cgc tca ctc ctt cga ggc ttc ctg ctg cag aac gag ttt gtg   1133
Cys Ala Arg Ser Leu Leu Arg Gly Phe Leu Leu Gln Asn Glu Phe Val
                320                 325                 330

ggg ttc atg tgg cgg cag cgg gga cgg gtc atc agc ctg tgg gag cgg   1181
Gly Phe Met Trp Arg Gln Arg Gly Arg Val Ile Ser Leu Trp Glu Arg
            335                 340                 345

ctg gaa ttt gtc aat ggc tgg tac atc ctg ctc gtc acc agc gat gtg   1229
Leu Glu Phe Val Asn Gly Trp Tyr Ile Leu Leu Val Thr Ser Asp Val
```

```
        350                 355                 360

ctc acc atc tcg ggc acc atc atg aag atc ggc atc gag gcc aag aac   1277
Leu Thr Ile Ser Gly Thr Ile Met Lys Ile Gly Ile Glu Ala Lys Asn
    365                 370                 375

ttg gcg agc tac gac gtc tgc agc atc ctc ctg ggc acc tcg acg ctg   1325
Leu Ala Ser Tyr Asp Val Cys Ser Ile Leu Leu Gly Thr Ser Thr Leu
380                 385                 390                 395

ctg gtg tgg gtg ggc gtg atc cgc tac ctg acc ttc ttc cac aac tac   1373
Leu Val Trp Val Gly Val Ile Arg Tyr Leu Thr Phe Phe His Asn Tyr
                400                 405                 410

aat atc ctc atc gcc aca ctg cgg gtg gcc ctg ccc agc gtc atg cgc   1421
Asn Ile Leu Ile Ala Thr Leu Arg Val Ala Leu Pro Ser Val Met Arg
            415                 420                 425

ttc tgc tgc tgc gtg gct gtc atc tac ctg ggc tac tgc ttc tgt ggc   1469
Phe Cys Cys Cys Val Ala Val Ile Tyr Leu Gly Tyr Cys Phe Cys Gly
        430                 435                 440

tgg atc gtg ctg ggg ccc tat cat gtg aag ttc cgc tca ctc tcc atg   1517
Trp Ile Val Leu Gly Pro Tyr His Val Lys Phe Arg Ser Leu Ser Met
    445                 450                 455

gtg tct gag tgc ctg ttc tcg ctc atc aat ggg gac gac atg ttt gtg   1565
Val Ser Glu Cys Leu Phe Ser Leu Ile Asn Gly Asp Asp Met Phe Val
460                 465                 470                 475

acg ttc gcc gcc atg cag gcg cag cag ggc cgc agc agc ctg gtg tgg   1613
Thr Phe Ala Ala Met Gln Ala Gln Gln Gly Arg Ser Ser Leu Val Trp
                480                 485                 490

ctc ttc tcc cag ctc tac ctt tac tcc ttc atc agc ctc ttc atc tac   1661
Leu Phe Ser Gln Leu Tyr Leu Tyr Ser Phe Ile Ser Leu Phe Ile Tyr
            495                 500                 505

atg gtg ctc agc ctc ttc atc gcg ctc atc acc ggc gcc tac gac acc   1709
Met Val Leu Ser Leu Phe Ile Ala Leu Ile Thr Gly Ala Tyr Asp Thr
        510                 515                 520

atc aag cat ccc ggc ggc gca ggc gca gag gag agc gag ctg cag gcc   1757
Ile Lys His Pro Gly Gly Ala Gly Ala Glu Glu Ser Glu Leu Gln Ala
    525                 530                 535

tac atc gca cag tgc cag gac agc ccc acc tcc ggc aag ttc cgc cgc   1805
Tyr Ile Ala Gln Cys Gln Asp Ser Pro Thr Ser Gly Lys Phe Arg Arg
540                 545                 550                 555

ggg agc ggc tcg gcc tgc agc ctt ctc tgc tgc tgc gga agg gac ccc   1853
Gly Ser Gly Ser Ala Cys Ser Leu Leu Cys Cys Cys Gly Arg Asp Pro
                560                 565                 570

tcg gag gag cat tcg ctg ctg gtg aat tgattcgacc tgactgccgt         1900
Ser Glu Glu His Ser Leu Leu Val Asn
            575                 580

tggaccgtag gccctggact gcagagaccc ccgccccgac cccgcttatt tatttgtagg 1960

gtttgctttt aaggatcggc tccctgtcgc gcccgaggag ggcctggacc tttcgtgtcg 2020
```

gacccttggg ggcggggaga ctgggtggga aggtgtgaat aaaagggaaa ttaaaaaaaa 2080

aaaaaaaatt ttaaa 2095

<210> 2
<211> 580
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Thr Ala Pro Ala Gly Pro Arg Gly Ser Glu Thr Glu Arg Leu Leu
  1               5                  10                  15

Thr Pro Asn Pro Gly Tyr Gly Thr Gln Ala Gly Pro Ser Pro Ala Pro
             20                  25                  30

Pro Thr Pro Pro Glu Glu Glu Asp Leu Arg Arg Arg Leu Lys Tyr Phe
         35                  40                  45

Phe Met Ser Pro Cys Asp Lys Phe Arg Ala Lys Gly Arg Lys Pro Cys
     50                  55                  60

Lys Leu Met Leu Gln Val Val Lys Ile Leu Val Val Thr Val Gln Leu
 65                  70                  75                  80

Ile Leu Phe Gly Leu Ser Asn Gln Leu Ala Val Thr Phe Arg Glu Glu
                 85                  90                  95

Asn Thr Ile Ala Phe Arg His Leu Phe Leu Leu Gly Tyr Ser Asp Gly
            100                 105                 110

Ala Asp Asp Thr Phe Ala Ala Tyr Thr Arg Glu Gln Leu Tyr Gln Ala
        115                 120                 125

Ile Phe His Ala Val Asp Gln Tyr Leu Ala Leu Pro Asp Val Ser Leu
    130                 135                 140

Gly Arg Tyr Ala Tyr Val Arg Gly Gly Gly Asp Pro Trp Thr Asn Gly
145                 150                 155                 160

Ser Gly Leu Ala Leu Cys Gln Arg Tyr Tyr His Arg Gly His Val Asp
                165                 170                 175

Pro Ala Asn Asp Thr Phe Asp Ile Asp Pro Met Val Val Thr Asp Cys
            180                 185                 190

Ile Gln Val Asp Pro Pro Glu Arg Pro Pro Pro Pro Pro Ser Asp Asp
        195                 200                 205

Leu Thr Leu Leu Glu Ser Ser Ser Ser Tyr Lys Asn Leu Thr Leu Lys
    210                 215                 220

Phe His Lys Leu Val Asn Val Thr Ile His Phe Arg Leu Lys Thr Ile
225                 230                 235                 240

Asn Leu Gln Ser Leu Ile Asn Asn Glu Ile Pro Asp Cys Tyr Thr Phe
                245                 250                 255

Ser Val Leu Ile Thr Phe Asp Asn Lys Ala His Ser Gly Arg Ile Pro
```

260 265 270

Ile Ser Leu Glu Thr Gln Ala His Ile Gln Glu Cys Lys His Pro Ser
275 280 285

Val Phe Gln His Gly Asp Asn Ser Phe Arg Leu Leu Phe Asp Val Val
290 295 300

Val Ile Leu Thr Cys Ser Leu Ser Phe Leu Leu Cys Ala Arg Ser Leu
305 310 315 320

Leu Arg Gly Phe Leu Leu Gln Asn Glu Phe Val Gly Phe Met Trp Arg
325 330 335

Gln Arg Gly Arg Val Ile Ser Leu Trp Glu Arg Leu Glu Phe Val Asn
340 345 350

Gly Trp Tyr Ile Leu Leu Val Thr Ser Asp Val Leu Thr Ile Ser Gly
355 360 365

Thr Ile Met Lys Ile Gly Ile Glu Ala Lys Asn Leu Ala Ser Tyr Asp
370 375 380

Val Cys Ser Ile Leu Leu Gly Thr Ser Thr Leu Leu Val Trp Val Gly
385 390 395 400

Val Ile Arg Tyr Leu Thr Phe Phe His Asn Tyr Asn Ile Leu Ile Ala
405 410 415

Thr Leu Arg Val Ala Leu Pro Ser Val Met Arg Phe Cys Cys Cys Val
420 425 430

Ala Val Ile Tyr Leu Gly Tyr Cys Phe Cys Gly Trp Ile Val Leu Gly
435 440 445

Pro Tyr His Val Lys Phe Arg Ser Leu Ser Met Val Ser Glu Cys Leu
450 455 460

Phe Ser Leu Ile Asn Gly Asp Asp Met Phe Val Thr Phe Ala Ala Met
465 470 475 480

Gln Ala Gln Gln Gly Arg Ser Ser Leu Val Trp Leu Phe Ser Gln Leu
485 490 495

Tyr Leu Tyr Ser Phe Ile Ser Leu Phe Ile Tyr Met Val Leu Ser Leu
500 505 510

Phe Ile Ala Leu Ile Thr Gly Ala Tyr Asp Thr Ile Lys His Pro Gly
515 520 525

Gly Ala Gly Ala Glu Glu Ser Glu Leu Gln Ala Tyr Ile Ala Gln Cys
530 535 540

Gln Asp Ser Pro Thr Ser Gly Lys Phe Arg Arg Gly Ser Gly Ser Ala
545 550 555 560

Cys Ser Leu Leu Cys Cys Cys Gly Arg Asp Pro Ser Glu Glu His Ser
565 570 575

Leu Leu Val Asn
580

<210> 3
<211> 1740
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1740)

<400> 3

```
atg aca gcc ccg gcg ggt ccg cgc ggc tca gag acc gag cgg ctt ctg   48
Met Thr Ala Pro Ala Gly Pro Arg Gly Ser Glu Thr Glu Arg Leu Leu
  1               5                  10                  15

acc ccc aac ccc ggg tat ggg acc cag gcg ggg cct tca ccg gcc cct   96
Thr Pro Asn Pro Gly Tyr Gly Thr Gln Ala Gly Pro Ser Pro Ala Pro
             20                  25                  30

ccg aca ccc cca gaa gag gaa gac ctt cgc cgt cgt ctc aaa tac ttt   144
Pro Thr Pro Pro Glu Glu Glu Asp Leu Arg Arg Arg Leu Lys Tyr Phe
         35                  40                  45

ttc atg agt ccc tgc gac aag ttt cga gcc aag ggc cgc aag ccc tgc   192
Phe Met Ser Pro Cys Asp Lys Phe Arg Ala Lys Gly Arg Lys Pro Cys
     50                  55                  60

aag ctg atg ctg caa gtg gtc aag atc ctg gtg gtc acg gtg cag ctc   240
Lys Leu Met Leu Gln Val Val Lys Ile Leu Val Val Thr Val Gln Leu
 65                  70                  75                  80

atc ctg ttt ggg ctc agt aat cag ctg gct gtg aca ttc cgg gaa gag   288
Ile Leu Phe Gly Leu Ser Asn Gln Leu Ala Val Thr Phe Arg Glu Glu
                 85                  90                  95

aac acc atc gcc ttc cga cac ctc ttc ctg ctg ggc tac tcg gac gga   336
Asn Thr Ile Ala Phe Arg His Leu Phe Leu Leu Gly Tyr Ser Asp Gly
            100                 105                 110

gcg gat gac acc ttc gca gcc tac acg cgg gag cag ctg tac cag gcc   384
Ala Asp Asp Thr Phe Ala Ala Tyr Thr Arg Glu Gln Leu Tyr Gln Ala
        115                 120                 125

atc ttc cat gct gtg gac cag tac ctg gcg ttg cct gac gtg tca ctg   432
Ile Phe His Ala Val Asp Gln Tyr Leu Ala Leu Pro Asp Val Ser Leu
    130                 135                 140

ggc cgg tat gcg tat gtc cgt ggt ggg ggt gac cct tgg acc aat ggc   480
Gly Arg Tyr Ala Tyr Val Arg Gly Gly Gly Asp Pro Trp Thr Asn Gly
145                 150                 155                 160

tca ggg ctt gct ctc tgc cag cgg tac tac cac cga ggc cac gtg gac   528
Ser Gly Leu Ala Leu Cys Gln Arg Tyr Tyr His Arg Gly His Val Asp
                165                 170                 175

ccg gcc aac gac aca ttt gac att gat ccg atg gtg gtt act gac tgc   576
Pro Ala Asn Asp Thr Phe Asp Ile Asp Pro Met Val Val Thr Asp Cys
            180                 185                 190

atc cag gtg gat ccc ccc gag cgg ccc cct ccg ccc ccc agc gac gat   624
```

```
Ile Gln Val Asp Pro Pro Glu Arg Pro Pro Pro Pro Ser Asp Asp
        195                 200                 205

ctc acc ctc ttg gaa agc agc tcc agt tac aag aac ctc acg ctc aaa   672
Leu Thr Leu Leu Glu Ser Ser Ser Ser Tyr Lys Asn Leu Thr Leu Lys
    210                 215                 220

ttc cac aag ctg gtc aat gtc acc atc cac ttc cgg ctg aag acc att   720
Phe His Lys Leu Val Asn Val Thr Ile His Phe Arg Leu Lys Thr Ile
225                 230                 235                 240

aac ctc cag agc ctc atc aat aat gag atc ccg gac tgc tat acc ttc   768
Asn Leu Gln Ser Leu Ile Asn Asn Glu Ile Pro Asp Cys Tyr Thr Phe
                245                 250                 255

agc gtc ctg atc acg ttt gac aac aaa gca cac agt ggg cgg atc ccc   816
Ser Val Leu Ile Thr Phe Asp Asn Lys Ala His Ser Gly Arg Ile Pro
            260                 265                 270

atc agc ctg gag acc cag gcc cac atc cag gag tgt aag cac ccc agt   864
Ile Ser Leu Glu Thr Gln Ala His Ile Gln Glu Cys Lys His Pro Ser
        275                 280                 285

gtc ttc cag cac gga gac aac agc ttc cgg ctc ctg ttt gac gtg gtg   912
Val Phe Gln His Gly Asp Asn Ser Phe Arg Leu Leu Phe Asp Val Val
    290                 295                 300

gtc atc ctc acc tgc tcc ctg tcc ttc ctc ctc tgc gcc cgc tca ctc   960
Val Ile Leu Thr Cys Ser Leu Ser Phe Leu Leu Cys Ala Arg Ser Leu
305                 310                 315                 320

ctt cga ggc ttc ctg ctg cag aac gag ttt gtg ggg ttc atg tgg cgg   1008
Leu Arg Gly Phe Leu Leu Gln Asn Glu Phe Val Gly Phe Met Trp Arg
                325                 330                 335

cag cgg gga cgg gtc atc agc ctg tgg gag cgg ctg gaa ttt gtc aat   1056
Gln Arg Gly Arg Val Ile Ser Leu Trp Glu Arg Leu Glu Phe Val Asn
            340                 345                 350

ggc tgg tac atc ctg ctc gtc acc agc gat gtg ctc acc atc tcg ggc   1104
Gly Trp Tyr Ile Leu Leu Val Thr Ser Asp Val Leu Thr Ile Ser Gly
        355                 360                 365

acc atc atg aag atc ggc atc gag gcc aag aac ttg gcg agc tac gac   1152
Thr Ile Met Lys Ile Gly Ile Glu Ala Lys Asn Leu Ala Ser Tyr Asp
    370                 375                 380

gtc tgc agc atc ctc ctg ggc acc tcg acg ctg ctg gtg tgg gtg ggc   1200
Val Cys Ser Ile Leu Leu Gly Thr Ser Thr Leu Leu Val Trp Val Gly
385                 390                 395                 400

gtg atc cgc tac ctg acc ttc ttc cac aac tac aat atc ctc atc gcc   1248
Val Ile Arg Tyr Leu Thr Phe Phe His Asn Tyr Asn Ile Leu Ile Ala
                405                 410                 415

aca ctg cgg gtg gcc ctg ccc agc gtc atg cgc ttc tgc tgc tgc gtg   1296
Thr Leu Arg Val Ala Leu Pro Ser Val Met Arg Phe Cys Cys Cys Val
            420                 425                 430

gct gtc atc tac ctg ggc tac tgc ttc tgt ggc tgg atc gtg ctg ggg   1344
Ala Val Ile Tyr Leu Gly Tyr Cys Phe Cys Gly Trp Ile Val Leu Gly
```

```
        435                 440                 445

ccc tat cat gtg aag ttc cgc tca ctc tcc atg gtg tct gag tgc ctg   1392
Pro Tyr His Val Lys Phe Arg Ser Leu Ser Met Val Ser Glu Cys Leu
    450                 455                 460

ttc tcg ctc atc aat ggg gac gac atg ttt gtg acg ttc gcc gcc atg   1440
Phe Ser Leu Ile Asn Gly Asp Asp Met Phe Val Thr Phe Ala Ala Met
465                 470                 475                 480

cag gcg cag cag ggc cgc agc agc ctg gtg tgg ctc ttc tcc cag ctc   1488
Gln Ala Gln Gln Gly Arg Ser Ser Leu Val Trp Leu Phe Ser Gln Leu
                485                 490                 495

tac ctt tac tcc ttc atc agc ctc ttc atc tac atg gtg ctc agc ctc   1536
Tyr Leu Tyr Ser Phe Ile Ser Leu Phe Ile Tyr Met Val Leu Ser Leu
            500                 505                 510

ttc atc gcg ctc atc acc ggc gcc tac gac acc atc aag cat ccc ggc   1584
Phe Ile Ala Leu Ile Thr Gly Ala Tyr Asp Thr Ile Lys His Pro Gly
        515                 520                 525

ggc gca ggc gca gag gag agc gag ctg cag gcc tac atc gca cag tgc   1632
Gly Ala Gly Ala Glu Glu Ser Glu Leu Gln Ala Tyr Ile Ala Gln Cys
    530                 535                 540

cag gac agc ccc acc tcc ggc aag ttc cgc cgc ggg agc ggc tcg gcc   1680
Gln Asp Ser Pro Thr Ser Gly Lys Phe Arg Arg Gly Ser Gly Ser Ala
545                 550                 555                 560

tgc agc ctt ctc tgc tgc tgc gga agg gac ccc tcg gag gag cat tcg   1728
Cys Ser Leu Leu Cys Cys Cys Gly Arg Asp Pro Ser Glu Glu His Ser
                565                 570                 575

ctg ctg gtg aat                                                  1740
Leu Leu Val Asn
            580


<210> 4
<211> 764
<212> PRT
<213> Homo sapiens

<400> 4
Met Thr Ser Pro Ser Ser Ser Pro Val Phe Arg Leu Glu Thr Leu Asp
  1               5                  10                  15

Gly Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg Gly Lys Leu Asp Phe
            20                  25                  30

Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe Gln Gly Glu Asp Arg
        35                  40                  45

Lys Phe Ala Pro Gln Ile Arg Val Asn Leu Asn Tyr Arg Lys Gly Thr
    50                  55                  60

Gly Ala Ser Gln Pro Asp Pro Asn Arg Phe Asp Arg Asp Arg Leu Phe
65                  70                  75                  80

Asn Ala Val Ser Arg Gly Val Pro Glu Asp Leu Ala Gly Leu Pro Glu
```

85 90 95

Tyr Leu Ser Lys Thr Ser Lys Tyr Leu Thr Asp Ser Glu Tyr Thr Glu
100 105 110

Gly Ser Thr Gly Lys Thr Cys Leu Met Lys Ala Val Leu Asn Leu Lys
115 120 125

Asp Gly Val Asn Ala Cys Ile Leu Pro Leu Leu Gln Ile Asp Arg Asp
130 135 140

Ser Gly Asn Pro Gln Pro Leu Val Asn Ala Gln Cys Thr Asp Asp Tyr
145 150 155 160

Tyr Arg Gly His Ser Ala Leu His Ile Ala Ile Glu Lys Arg Ser Leu
165 170 175

Gln Cys Val Lys Leu Leu Val Glu Asn Gly Ala Asn Val His Ala Arg
180 185 190

Ala Cys Gly Arg Phe Phe Gln Lys Gly Gln Gly Thr Cys Phe Tyr Phe
195 200 205

Gly Glu Leu Pro Leu Ser Leu Ala Ala Cys Thr Lys Gln Trp Asp Val
210 215 220

Val Ser Tyr Leu Leu Glu Asn Pro His Gln Pro Ala Ser Leu Gln Ala
225 230 235 240

Thr Asp Ser Gln Gly Asn Thr Val Leu His Ala Leu Val Met Ile Ser
245 250 255

Asp Asn Ser Ala Glu Asn Ile Ala Leu Val Thr Ser Met Tyr Asp Gly
260 265 270

Leu Leu Gln Ala Gly Ala Arg Leu Cys Pro Thr Val Gln Leu Glu Asp
275 280 285

Ile Arg Asn Leu Gln Asp Leu Thr Pro Leu Lys Leu Ala Ala Lys Glu
290 295 300

Gly Lys Ile Glu Ile Phe Arg His Ile Leu Gln Arg Glu Phe Ser Gly
305 310 315 320

Leu Ser His Leu Ser Arg Lys Phe Thr Glu Trp Cys Tyr Gly Pro Val
325 330 335

Arg Val Ser Leu Tyr Asp Leu Ala Ser Val Asp Ser Cys Glu Glu Asn
340 345 350

Ser Val Leu Glu Ile Ile Ala Phe His Cys Lys Ser Pro His Arg His
355 360 365

Arg Met Val Val Leu Glu Pro Leu Asn Lys Leu Leu Gln Ala Lys Trp
370 375 380

Asp Leu Leu Ile Pro Lys Phe Phe Leu Asn Phe Leu Cys Asn Leu Ile
385 390 395 400

Tyr Met Phe Ile Phe Thr Ala Val Ala Tyr His Gln Pro Thr Leu Lys
405 410 415

```
Lys Gln Ala Ala Pro His Leu Lys Ala Glu Val Gly Asn Ser Met Leu
            420                 425                 430

Leu Thr Gly His Ile Leu Ile Leu Leu Gly Gly Ile Tyr Leu Leu Val
        435                 440                 445

Gly Gln Leu Trp Tyr Phe Trp Arg Arg His Val Phe Ile Trp Ile Ser
    450                 455                 460

Phe Ile Asp Ser Tyr Phe Glu Ile Leu Phe Leu Phe Gln Ala Leu Leu
465                 470                 475                 480

Thr Val Val Ser Gln Val Leu Cys Phe Leu Ala Ile Glu Trp Tyr Leu
                485                 490                 495

Pro Leu Leu Val Ser Ala Leu Val Leu Gly Trp Leu Asn Leu Leu Tyr
            500                 505                 510

Tyr Thr Arg Gly Phe Gln His Thr Gly Ile Tyr Ser Val Met Ile Gln
        515                 520                 525

Lys Val Ile Leu Arg Asp Leu Leu Arg Phe Leu Leu Ile Tyr Leu Val
    530                 535                 540

Phe Leu Phe Gly Phe Ala Val Ala Leu Val Ser Leu Ser Gln Glu Ala
545                 550                 555                 560

Trp Arg Pro Glu Ala Pro Thr Gly Pro Asn Ala Thr Glu Ser Val Gln
                565                 570                 575

Pro Met Glu Gly Gln Glu Asp Glu Gly Asn Gly Ala Gln Tyr Arg Gly
            580                 585                 590

Ile Leu Glu Ala Ser Leu Glu Leu Phe Lys Phe Thr Ile Gly Met Gly
        595                 600                 605

Glu Leu Ala Phe Gln Glu Gln Leu His Phe Arg Gly Met Val Leu Leu
    610                 615                 620

Leu Leu Leu Ala Tyr Val Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met
625                 630                 635                 640

Leu Ile Ala Leu Met Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser
                645                 650                 655

Trp Ser Ile Trp Lys Leu Gln Lys Ala Ile Ser Val Leu Glu Met Glu
            660                 665                 670

Asn Gly Tyr Trp Trp Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu
        675                 680                 685

Thr Val Gly Thr Lys Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe
    690                 695                 700

Arg Val Glu Glu Val Asn Trp Ala Ser Trp Glu Gln Thr Leu Pro Thr
705                 710                 715                 720

Leu Cys Glu Asp Pro Ser Gly Ala Gly Val Pro Arg Thr Leu Glu Asn
                725                 730                 735
```

```
Pro Val Leu Ala Ser Pro Pro Lys Glu Asp Glu Asp Gly Ala Ser Glu
            740                 745                 750

Glu Asn Tyr Val Pro Val Gln Leu Leu Gln Ser Asn
        755                 760
```

<210> 5
<211> 764
<212> PRT
<213> Homo sapiens

<220>
<223> Xaa at position 667 may be any amino acid

<400> 5

```
Met Thr Ser Pro Ser Ser Ser Pro Val Phe Arg Leu Glu Thr Leu Asp
  1               5                  10                  15

Gly Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg Gly Lys Leu Asp Phe
             20                  25                  30

Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe Gln Gly Glu Asp Arg
         35                  40                  45

Lys Phe Ala Ser Gln Ile Arg Val Asn Leu Asn Tyr Arg Lys Gly Thr
     50                  55                  60

Gly Ala Ser Gln Pro Asp Pro Asn Arg Phe Asp Arg Asp Arg Leu Phe
 65                  70                  75                  80

Asn Val Val Ser Arg Gly Val Pro Glu Asp Leu Ala Gly Leu Pro Glu
                 85                  90                  95

Tyr Leu Ser Lys Thr Ser Lys Tyr Leu Thr Asp Ser Glu Tyr Thr Glu
            100                 105                 110

Gly Ser Thr Gly Lys Thr Cys Leu Met Lys Ala Val Leu Asn Leu Lys
        115                 120                 125

Asp Gly Val Asn Ala Cys Ile Leu Pro Leu Leu Gln Ile Asp Arg Asp
    130                 135                 140

Ser Gly Asn Pro Gln Pro Leu Val Asn Ala Gln Cys Thr Asp Asp Tyr
145                 150                 155                 160

Tyr Arg Gly His Ser Ala Leu His Ile Ala Ile Glu Lys Arg Ser Leu
                165                 170                 175

Gln Cys Val Lys Leu Leu Val Glu Asn Gly Ala Asn Val His Ala Arg
            180                 185                 190

Ala Cys Gly Arg Phe Phe Gln Asn Gly Gln Gly Thr Cys Phe Tyr Phe
        195                 200                 205

Gly Glu Leu Pro Leu Ser Leu Ala Ala Cys Thr Lys Gln Trp Asp Val
    210                 215                 220

Val Ser Tyr Leu Leu Glu Asn Pro His Gln Pro Ala Ser Leu Gln Ala
225                 230                 235                 240
```

```
Thr Asp Ser Gln Gly Asn Thr Val Leu His Ala Leu Val Met Ile Ser
                245                 250                 255

Asp Asn Ser Ala Glu Asn Ile Ala Leu Val Thr Ser Met Tyr Asp Gly
            260                 265                 270

Leu Leu Gln Ala Gly Ala Arg Leu Cys Pro Thr Val Gln Leu Glu Asp
        275                 280                 285

Ile Arg Asn Leu Gln Asp Leu Thr Pro Leu Lys Leu Ala Ala Lys Glu
    290                 295                 300

Gly Lys Ile Glu Ile Phe Arg His Ile Leu Gln Arg Glu Phe Ser Gly
305                 310                 315                 320

Leu Ser His Leu Ser Arg Lys Phe Thr Glu Trp Cys Tyr Gly Pro Val
                325                 330                 335

Arg Val Ser Leu Tyr Asp Leu Ala Ser Val Asp Ser Cys Glu Glu Asn
            340                 345                 350

Ser Val Leu Glu Ile Ile Ala Phe His Cys Lys Ser Pro His Arg His
        355                 360                 365

Arg Met Val Val Leu Glu Pro Leu Asn Lys Leu Leu Gln Ala Lys Trp
    370                 375                 380

Asp Leu Leu Ile Pro Lys Phe Phe Leu Asn Phe Leu Cys Asn Leu Ile
385                 390                 395                 400

Tyr Met Phe Ile Phe Thr Ala Val Ala Tyr His Gln Pro Thr Leu Lys
                405                 410                 415

Lys Gln Ala Ala Pro His Leu Lys Ala Glu Val Gly Asn Ser Met Leu
            420                 425                 430

Leu Thr Gly His Ile Leu Ile Leu Leu Gly Gly Ile Tyr Leu Leu Val
        435                 440                 445

Gly Gln Leu Trp Tyr Phe Trp Arg Arg His Leu Phe Ile Trp Ile Ser
    450                 455                 460

Tyr Thr Asp Ser Tyr Phe Glu Ile Leu Phe Leu Phe His Ser Leu Leu
465                 470                 475                 480

Thr Val Val Ser Leu Val Leu Cys Phe Leu Val Ile Glu Trp Tyr Leu
                485                 490                 495

Pro Leu Leu Val Ser Ala Leu Val Leu Gly Trp Leu Asn Leu Leu Tyr
            500                 505                 510

Tyr Thr Arg Gly Phe Gln His Thr Gly Ile Tyr Ser Val Met Ile Gln
        515                 520                 525

Lys Val Ile Leu Arg Asp Met Val Arg Phe Leu Val Ile Tyr Leu Val
    530                 535                 540

Phe Leu Phe Gly Phe Ala Val Ala Leu Val Ser Leu Ser Gln Glu Ala
545                 550                 555                 560

Trp Arg Pro Glu Ala Pro Thr Gly Pro Asn Ala Thr Glu Ser Val Gln
```

```
                565                 570                 575

Pro Met Glu Gly Gln Glu Asp Glu Gly Asn Gly Ala Gln Tyr Arg Gly
            580                 585                 590

Ile Leu Glu Ala Ser Leu Glu Leu Phe Lys Phe Thr Ile Gly Met Gly
        595                 600                 605

Glu Leu Ala Phe Gln Glu Gln Leu His Phe Arg Gly Met Val Leu Leu
    610                 615                 620

Leu Leu Leu Ala Tyr Val Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met
625                 630                 635                 640

Leu Ile Ala Leu Met Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser
                645                 650                 655

Trp Ser Ile Trp Lys Leu Gln Lys Ala Ile Xaa Val Leu Glu Met Glu
            660                 665                 670

Asn Gly Tyr Trp Trp Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu
        675                 680                 685

Thr Val Gly Thr Lys Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe
    690                 695                 700

Arg Val Glu Glu Val Asn Trp Ala Ser Trp Glu Gln Thr Leu Pro Thr
705                 710                 715                 720

Leu Cys Glu Asp Pro Ser Gly Ala Gly Val Pro Arg Thr Leu Glu Asn
                725                 730                 735

Pro Val Leu Ala Ser Pro Pro Lys Glu Asp Glu Asp Gly Ala Ser Glu
            740                 745                 750

Glu Asn Tyr Val Pro Val Gln Leu Leu Gln Ser Asn
        755                 760


<210> 6
<211> 966
<212> PRT
<213> Homo sapiens

<400> 6
Met Val Asn Ser Arg Arg Val Gln Pro Gln Pro Pro Gly Asp Ala Gly
  1               5                  10                  15

Arg Ser Pro Ala Pro Arg Ala Ser Gly Pro Gly Arg Leu Val Ala Gly
            20                  25                  30

Gly Ala Gly Leu Ala Val Pro Gly Gly Leu Gly Glu Gln Arg Gly Leu
        35                  40                  45

Glu Ile Glu Met Glu Arg Ile Arg Gln Ala Ala Ala Arg Asp Pro Pro
    50                  55                  60

Ala Gly Ala Ser Ala Ser Pro Ser Pro Pro Leu Ser Ser Cys Ser Arg
65                  70                  75                  80

Gln Ala Trp Ser Arg Asp Asn Pro Gly Phe Glu Ala Glu Glu Asp Asp
```

85 90 95

Asp Asp Asp Glu Val Glu Gly Glu Glu Gly Gly Met Val Val Glu Met
100 105 110

Asp Val Glu Trp Arg Pro Gly Ser Arg Arg Ser Ala Ser Ser Ser Ala
115 120 125

Val Ser Ser Val Gly Ala Arg Gly Arg Gly Leu Gly Ser Tyr Arg Gly
130 135 140

Ala Ala His Leu Ser Gly Arg Arg Arg Arg Leu Glu Asp Gln Gly Ala
145 150 155 160

Gln Cys Pro Ser Pro Ala Gly Gly Gly Asp Pro Leu His Arg His Leu
165 170 175

Pro Leu Glu Gly Gln Pro Pro Arg Val Ala Trp Ala Glu Arg Leu Val
180 185 190

Arg Gly Leu Arg Gly Leu Trp Gly Thr Arg Leu Met Glu Glu Ser Asn
195 200 205

Ala Asn Arg Glu Lys Tyr Leu Lys Ser Val Leu Arg Glu Leu Val Thr
210 215 220

Tyr Leu Phe Phe Leu Val Val Leu Cys Ile Leu Thr Tyr Gly Met Met
225 230 235 240

Ser Ser Asn Val Tyr Tyr Tyr Thr Arg Thr Leu Ser Gln Leu Phe Ile
245 250 255

Asp Thr Pro Val Ser Lys Thr Glu Lys Thr Asn Phe Lys Thr Leu Ser
260 265 270

Ser Met Glu Asp Phe Trp Lys Phe Thr Glu Gly Ser Phe Leu Asp Gly
275 280 285

Leu Tyr Trp Lys Ala Gln Thr Ser Asn His Thr Gln Ala Asp Asn Arg
290 295 300

Ser Phe Ile Phe Tyr Glu Asn Leu Leu Leu Gly Val Pro Arg Leu Arg
305 310 315 320

Gln Leu Arg Val Arg Asn Gly Ser Cys Ser Ile Pro Gln Asp Leu Arg
325 330 335

Asp Glu Ile Lys Glu Cys Tyr Asp Val Tyr Ser Val Ser Ser Glu Asp
340 345 350

Arg Ala Pro Phe Gly Pro Arg Asn Gly Thr Ala Trp Met Tyr Thr Ser
355 360 365

Glu Lys Glu Leu Asn Gly Ser Ser His Trp Gly Ile Ile Ala Ser Tyr
370 375 380

Ser Gly Ala Gly Tyr Tyr Leu Asp Leu Ser Arg Thr Arg Glu Glu Thr
385 390 395 400

Ala Ala Gln Leu Ala Gly Leu Arg Arg Asn Phe Trp Leu Asp Arg Gly
405 410 415

```
Thr Arg Ala Ala Phe Ile Asp Phe Ser Val Tyr Asn Ala Asn Ile Asn
            420                 425                 430

Leu Phe Cys Val Val Arg Leu Leu Ala Glu Phe Pro Ala Thr Gly Gly
        435                 440                 445

Val Val Pro Ser Trp Gln Phe Gln Pro Val Lys Leu Ile Arg Tyr Val
    450                 455                 460

Thr Ala Phe Asp Phe Phe Leu Ala Ala Cys Glu Ile Ile Phe Cys Phe
465                 470                 475                 480

Phe Ile Ile Tyr Tyr Val Val Glu Glu Ile Leu Glu Ile Arg Ile His
                485                 490                 495

Arg Leu Ser Tyr Phe Arg Ser Phe Trp Asn Cys Leu Asp Val Val Ile
            500                 505                 510

Val Val Leu Ser Val Val Ala Met Val Ile Asn Ile Tyr Arg Met Ser
        515                 520                 525

Asn Ala Glu Gly Leu Leu Gln Phe Leu Glu Asp Gln Asn Ser Phe Pro
    530                 535                 540

Asn Phe Glu His Val Ala Tyr Trp Gln Ile Gln Phe Asn Asn Ile Ser
545                 550                 555                 560

Ala Val Met Val Phe Leu Val Trp Ile Lys Leu Phe Lys Phe Ile Asn
                565                 570                 575

Phe Asn Arg Thr Met Ser Gln Leu Ser Thr Thr Met Ser Arg Cys Ala
            580                 585                 590

Lys Asp Leu Phe Gly Phe Thr Ile Met Phe Ser Ile Ile Phe Leu Ala
        595                 600                 605

Tyr Ala Gln Leu Ala Tyr Leu Val Phe Gly Thr Gln Val Asp Asp Phe
    610                 615                 620

Ser Thr Phe Gln Glu Cys Ile Phe Thr Gln Phe Arg Ile Ile Leu Gly
625                 630                 635                 640

Asp Ile Asn Phe Ala Glu Ile Glu Glu Ala Asn Arg Val Leu Gly Pro
                645                 650                 655

Leu Tyr Phe Thr Thr Phe Val Phe Phe Met Phe Phe Ile Leu Leu Asn
            660                 665                 670

Met Phe Leu Ala Ile Ile Asn Asp Ser Tyr Ser Glu Val Lys Ser Ser
        675                 680                 685

Val Ala Gln Gln Lys Ala Glu Met Glu Leu Ser Asp Leu Ile Arg Lys
    690                 695                 700

Gly Cys Gln Lys Ala Leu Val Lys Leu Lys Leu Lys Arg Asn Thr Val
705                 710                 715                 720

Asp Ala Ile Ser Glu Ser Leu Arg Gln Gly Gly Gly Lys Leu Asn Phe
                725                 730                 735
```

```
Asp Glu Leu Arg Gln Asp Leu Lys Gly Glu Gly His Thr Asp Ala Glu
            740                 745                 750

Ile Glu Ala Ile Phe Thr Lys Tyr Asp Gln Asp Gly Asp Gln Glu Leu
        755                 760                 765

Thr Glu Arg Glu His Gln Gln Met Arg Asp Asp Leu Glu Lys Glu Arg
    770                 775                 780

Glu Asp Leu Asp Leu Glu His Ser Ser Leu Pro Arg Pro Met Ser Ser
785                 790                 795                 800

Arg Ser Phe Pro Arg Ser Leu Asp Asp Ser Glu Glu Glu Asp Asp Glu
                805                 810                 815

Asp Ser Gly His Ser Ser Arg Arg Arg Gly Ser Ile Ser Ser Gly Val
            820                 825                 830

Ser Tyr Glu Glu Phe Gln Val Leu Val Arg Arg Val Asp Arg Met Glu
        835                 840                 845

His Ser Ile Gly Ser Ile Val Ser Lys Ile Asp Ala Val Ile Val Lys
    850                 855                 860

Leu Glu Ile Met Glu Arg Ala Lys Leu Lys Arg Arg Glu Val Leu Gly
865                 870                 875                 880

Arg Leu Leu Asp Gly Val Ala Glu Asp Ala Arg Leu Gly Arg Asp Ser
                885                 890                 895

Glu Ile His Arg Glu Gln Met Glu Arg Leu Val Arg Glu Glu Leu Glu
            900                 905                 910

Arg Trp Glu Ser Asp Asp Ala Ala Ser Gln Thr Gly His Gly Val Ser
        915                 920                 925

Thr Gln Val Gly Leu Gly Gly Gln Pro His Pro Arg Asn Pro Arg Pro
    930                 935                 940

Pro Ser Ser Gln Ser Ala Glu Gly Leu Glu Gly Gly Gly Gly Asn Gly
945                 950                 955                 960

Ser Ala Asn Val His Ala
                965


<210> 7
<211> 1533
<212> PRT
<213> Homo sapiens

<400> 7
Met Tyr Ile Arg Val Ser Tyr Asp Thr Lys Pro Asp Ser Leu Leu His
  1               5                  10                  15

Leu Met Val Lys Asp Trp Gln Leu Glu Leu Pro Lys Leu Leu Ile Ser
            20                  25                  30

Val His Gly Gly Leu Gln Asn Phe Glu Met Gln Pro Lys Leu Lys Gln
         35                  40                  45
```

```
Val Phe Gly Lys Gly Leu Ile Lys Ala Ala Met Thr Thr Gly Ala Trp
     50                  55                  60

Ile Phe Thr Gly Gly Val Ser Thr Gly Val Ile Ser His Val Gly Asp
 65                  70                  75                  80

Ala Leu Lys Asp His Ser Ser Lys Ser Arg Gly Arg Val Cys Ala Ile
                 85                  90                  95

Gly Ile Ala Pro Trp Gly Ile Val Glu Asn Lys Glu Asp Leu Val Gly
            100                 105                 110

Lys Asp Val Thr Arg Val Tyr Gln Thr Met Ser Asn Pro Leu Ser Lys
        115                 120                 125

Leu Ser Val Leu Asn Asn Ser His Thr His Phe Ile Leu Ala Asp Asn
    130                 135                 140

Gly Thr Leu Gly Lys Tyr Gly Ala Glu Val Lys Leu Arg Arg Leu Leu
145                 150                 155                 160

Glu Lys His Ile Ser Leu Gln Lys Ile Asn Thr Arg Leu Gly Gln Gly
                165                 170                 175

Val Pro Leu Val Gly Leu Val Val Glu Gly Gly Pro Asn Val Val Ser
            180                 185                 190

Ile Val Leu Glu Tyr Leu Gln Glu Glu Pro Pro Ile Pro Val Val Ile
        195                 200                 205

Cys Asp Gly Ser Gly Arg Ala Ser Asp Ile Leu Ser Phe Ala His Lys
    210                 215                 220

Tyr Cys Glu Glu Gly Gly Ile Ile Asn Glu Ser Leu Arg Glu Gln Leu
225                 230                 235                 240

Leu Val Thr Ile Gln Lys Thr Phe Asn Tyr Asn Lys Ala Gln Ser His
                245                 250                 255

Gln Leu Phe Ala Ile Ile Met Glu Cys Met Lys Lys Lys Glu Leu Val
            260                 265                 270

Thr Val Phe Arg Met Gly Ser Glu Gly Gln Gln Asp Ile Glu Met Ala
        275                 280                 285

Ile Leu Thr Ala Leu Leu Lys Gly Thr Asn Val Ser Ala Pro Asp Gln
    290                 295                 300

Leu Ser Leu Ala Leu Ala Trp Asn Arg Val Asp Ile Ala Arg Ser Gln
305                 310                 315                 320

Ile Phe Val Phe Gly Pro His Trp Thr Pro Leu Gly Ser Leu Ala Pro
                325                 330                 335

Pro Thr Asp Ser Lys Ala Thr Glu Lys Glu Lys Lys Pro Pro Met Ala
            340                 345                 350

Thr Thr Lys Gly Gly Arg Gly Lys Gly Lys Gly Lys Lys Lys Gly Lys
        355                 360                 365

Val Lys Glu Glu Val Glu Glu Glu Thr Asp Pro Arg Lys Ile Glu Leu
```

370 375 380

Leu Asn Trp Val Asn Ala Leu Glu Gln Ala Met Leu Asp Ala Leu Val
385 390 395 400

Leu Asp Arg Val Asp Phe Val Lys Leu Leu Ile Glu Asn Gly Val Asn
405 410 415

Met Gln His Phe Leu Thr Ile Pro Arg Leu Glu Glu Leu Tyr Asn Thr
420 425 430

Arg Leu Gly Pro Pro Asn Thr Leu His Leu Leu Val Arg Asp Val Lys
435 440 445

Lys Ser Asn Leu Pro Pro Asp Tyr His Ile Ser Leu Ile Asp Ile Gly
450 455 460

Leu Val Leu Glu Tyr Leu Met Gly Gly Ala Tyr Arg Cys Asn Tyr Thr
465 470 475 480

Arg Lys Asn Phe Arg Thr Leu Tyr Asn Asn Leu Phe Gly Pro Lys Arg
485 490 495

Pro Lys Ala Leu Lys Leu Leu Gly Met Glu Asp Asp Glu Pro Pro Ala
500 505 510

Lys Gly Lys Lys Lys Lys Lys Lys Lys Lys Glu Glu Glu Ile Asp Ile
515 520 525

Asp Val Asp Asp Pro Ala Val Ser Arg Phe Gln Tyr Pro Phe His Glu
530 535 540

Leu Met Val Trp Ala Val Leu Met Lys Arg Gln Lys Met Ala Val Phe
545 550 555 560

Leu Trp Gln Arg Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys
565 570 575

Lys Leu Tyr Lys Ala Met Ala His Glu Ser Ser Glu Ser Asp Leu Val
580 585 590

Asp Asp Ile Ser Gln Asp Leu Asp Asn Asn Ser Lys Asp Phe Gly Gln
595 600 605

Leu Ala Leu Glu Leu Leu Asp Gln Ser Tyr Lys His Asp Glu Gln Ile
610 615 620

Ala Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr
625 630 635 640

Cys Leu Lys Leu Ala Val Ala Ala Lys His Arg Asp Phe Ile Ala His
645 650 655

Thr Cys Ser Gln Met Leu Leu Thr Asp Met Trp Met Gly Arg Leu Arg
660 665 670

Met Arg Lys Asn Pro Gly Leu Lys Val Ile Met Gly Ile Leu Leu Pro
675 680 685

Pro Thr Ile Leu Phe Leu Glu Phe Arg Thr Tyr Asp Asp Phe Ser Tyr
690 695 700

```
Gln Thr Ser Lys Glu Asn Glu Asp Gly Lys Glu Lys Glu Glu Glu Asn
705                 710                 715                 720

Thr Asp Ala Asn Ala Asp Ala Gly Ser Arg Lys Gly Asp Glu Glu Asn
                725                 730                 735

Glu His Lys Lys Gln Arg Ser Ile Pro Ile Gly Thr Lys Ile Cys Glu
            740                 745                 750

Phe Tyr Asn Ala Pro Ile Val Lys Phe Trp Phe Tyr Thr Ile Ser Tyr
        755                 760                 765

Leu Gly Tyr Leu Leu Leu Phe Asn Tyr Val Ile Leu Val Arg Met Asp
    770                 775                 780

Gly Trp Pro Ser Leu Gln Glu Trp Ile Val Ile Ser Tyr Ile Val Ser
785                 790                 795                 800

Leu Ala Leu Glu Lys Ile Arg Glu Ile Leu Met Ser Glu Pro Gly Lys
                805                 810                 815

Leu Ser Gln Lys Ile Lys Val Trp Leu Gln Glu Tyr Trp Asn Ile Thr
            820                 825                 830

Asp Leu Val Ala Ile Ser Thr Phe Met Ile Gly Ala Ile Leu Arg Leu
        835                 840                 845

Gln Asn Gln Pro Tyr Met Gly Tyr Gly Arg Val Ile Tyr Cys Val Asp
    850                 855                 860

Ile Ile Phe Trp Tyr Ile Arg Val Leu Asp Ile Phe Gly Val Asn Lys
865                 870                 875                 880

Tyr Leu Gly Pro Tyr Val Met Met Ile Gly Lys Met Met Ile Asp Met
                885                 890                 895

Leu Tyr Phe Val Val Ile Met Leu Val Val Leu Met Ser Phe Gly Val
            900                 905                 910

Ala Arg Gln Ala Ile Leu His Pro Glu Glu Lys Pro Ser Trp Lys Leu
        915                 920                 925

Ala Arg Asn Ile Phe Tyr Met Pro Tyr Trp Met Ile Tyr Gly Glu Val
    930                 935                 940

Phe Ala Asp Gln Ile Asp Leu Tyr Ala Met Glu Ile Asn Pro Pro Cys
945                 950                 955                 960

Gly Glu Asn Leu Tyr Asp Glu Glu Gly Lys Arg Leu Pro Pro Cys Ile
                965                 970                 975

Pro Gly Ala Trp Leu Thr Pro Ala Leu Met Ala Cys Tyr Leu Leu Val
            980                 985                 990

Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Val Phe Asn Asn Thr
        995                1000                1005

Phe Phe Glu Val Lys Ser Ile Ser Asn Gln Val Trp Lys Phe Gln Arg
   1010                1015                1020
```

```
Tyr Gln Leu Ile Met Thr Phe His Asp Arg Pro Val Leu Pro Pro Pro
1025                1030                1035                1040

Met Ile Ile Leu Ser His Ile Tyr Ile Ile Ile Met Arg Leu Ser Gly
                1045                1050                1055

Arg Cys Arg Lys Lys Arg Glu Gly Asp Gln Glu Glu Arg Asp Arg Gly
            1060                1065                1070

Leu Lys Leu Phe Leu Ser Asp Glu Glu Leu Lys Arg Leu His Glu Phe
        1075                1080                1085

Glu Glu Gln Cys Val Gln Glu His Phe Arg Glu Lys Glu Asp Glu Gln
    1090                1095                1100

Gln Ser Ser Ser Asp Glu Arg Ile Arg Val Thr Ser Glu Arg Val Glu
1105                1110                1115                1120

Asn Met Ser Met Arg Leu Glu Glu Ile Asn Glu Arg Glu Thr Phe Met
                1125                1130                1135

Lys Thr Ser Leu Gln Thr Val Asp Leu Arg Leu Ala Gln Leu Glu Glu
            1140                1145                1150

Leu Ser Asn Arg Met Val Asn Ala Leu Glu Asn Leu Ala Gly Ile Asp
        1155                1160                1165

Arg Ser Asp Leu Ile Gln Ala Arg Ser Arg Ala Ser Ser Glu Cys Glu
    1170                1175                1180

Ala Thr Tyr Leu Leu Arg Gln Ser Ser Ile Asn Ser Ala Asp Gly Tyr
1185                1190                1195                1200

Ser Leu Tyr Arg Tyr His Phe Asn Gly Glu Glu Leu Leu Phe Glu Asp
                1205                1210                1215

Thr Ser Leu Ser Thr Ser Pro Gly Thr Gly Val Arg Lys Lys Thr Cys
            1220                1225                1230

Ser Phe Arg Ile Lys Glu Glu Lys Asp Val Lys Thr His Leu Val Pro
        1235                1240                1245

Glu Cys Gln Asn Ser Leu His Leu Ser Leu Gly Thr Ser Thr Ser Ala
    1250                1255                1260

Thr Pro Asp Gly Ser His Leu Ala Val Asp Asp Leu Lys Asn Ala Glu
1265                1270                1275                1280

Glu Ser Lys Leu Gly Pro Asp Ile Gly Ile Ser Lys Glu Asp Asp Glu
                1285                1290                1295

Arg Gln Thr Asp Ser Lys Lys Glu Glu Thr Ile Ser Pro Ser Leu Asn
            1300                1305                1310

Lys Thr Asp Val Ile His Gly Gln Asp Lys Ser Asp Val Gln Asn Thr
        1315                1320                1325

Gln Leu Thr Val Glu Thr Thr Asn Ile Glu Gly Thr Ile Ser Tyr Pro
    1330                1335                1340

Leu Glu Glu Thr Lys Ile Thr Arg Tyr Phe Pro Asp Glu Thr Ile Asn
```

```
1345                1350                1355                1360

Ala Cys Lys Thr Met Lys Ser Arg Ser Phe Val Tyr Ser Arg Gly Arg
                1365                1370                1375

Lys Leu Val Gly Gly Val Asn Gln Asp Val Glu Tyr Ser Ser Ile Thr
            1380                1385                1390

Asp Gln Gln Leu Thr Thr Glu Trp Gln Cys Gln Val Gln Lys Ile Thr
        1395                1400                1405

Arg Ser His Ser Thr Asp Ile Pro Tyr Ile Val Ser Glu Ala Ala Val
    1410                1415                1420

Gln Ala Glu Gln Lys Glu Gln Phe Ala Asp Met Gln Asp Glu His His
1425                1430                1435                1440

Val Ala Glu Ala Ile Pro Arg Ile Pro Arg Leu Ser Leu Thr Ile Thr
                1445                1450                1455

Asp Arg Asn Gly Met Glu Asn Leu Leu Ser Val Lys Pro Asp Gln Thr
            1460                1465                1470

Leu Gly Phe Pro Ser Leu Arg Ser Lys Ser Leu His Gly His Pro Arg
        1475                1480                1485

Asn Val Lys Ser Ile Gln Gly Lys Leu Asp Arg Ser Gly His Ala Ser
    1490                1495                1500

Ser Val Ser Ser Leu Val Ile Val Ser Gly Met Thr Ala Glu Glu Lys
1505                1510                1515                1520

Lys Val Lys Lys Glu Lys Ala Ser Thr Glu Thr Glu Cys
                1525                1530
```

INTERNATIONAL SEARCH REPORT

Internatio pplication No
PCT/US 01/11442

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K16/18 C12Q1/68 G01N33/53
A61K38/17 A61K39/395 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 00 17222 A (HUAMN GENOME SCIENCES, INC.) 30 March 2000 (2000-03-30)	1-15
A	abstract page 2, line 7 - line 14 gene ID NO: 16 page 48, line 1 -page 51, line 26 page 92; table 1 page 94, line 1 -page 156, line 21 page 186, line 29 -page 210, line 23 page 311 -page 314; claims 1-16,18-21 SEQ ID NO: 26 figures PP14-15 SEQ ID NO: 63 figures PP44-46	16-30

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	*Date of mailing of the international search report*
28 August 2001	12/09/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Fuchs, U

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL, HEIDELBERG, FRG 'Online! 11 December 1998 (1998-12-11) NCI-CGAP: "tb24a12.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE: 2055262 3' similar to WP: R13A5.1 CE01370, mRNA sequence" Database accession no. AI307240 XP002176007 cited in the application	1-9
A	the whole document	10-15
X	DATABASE EMBL, HEIDELBERG, FRG 'Online! 12 July 1999 (1999-07-12) HILLIER, L. ET AL.: "au44h03.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone IMAGE: 2517653 3' similar to WP: R13A5.1 CE01370, mRNA sequence" Database accession no. AI816064 XP002176008 cited in the application	1-9
A	the whole document	10-15
X	-& DATABASE EMBL, HEIDELBERG, FRG 'Online! 12 July 1999 (1999-07-12) HILLIER, L. ET AL.: "au44h03.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone IMAGE: 2517653 5' similar to WP: R13A5.1 CE01370, mRNA sequence" Database accession no. AI815981 XP002176009	1-9
A	the whole document	10-15
P,X	WO 01 12662 A (INCYTE GENOMICS, INC.) 22 February 2001 (2001-02-22) abstract page 5, line 9 -page 10, line 16 page 23, line 32 -page 58, line 24 clone ID NO: 977658 page 74; table 1 SEQ ID NO: 13 page 81; table 2 SEQ ID NO: 50 page 91; table 3 page 107 -page 111; claims 1-27 SEQ ID NO: 13 figures PP12-13 SEQ ID NO: 50 figures PP44-45	1-15

-/--

INTERNATIONAL SEARCH REPORT

Internatic .pplication No
PCT/US 01/11442

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	SUN, M. ET AL.: "Mucolipidosis type IV is caused by mutations in a gene encoding a novel transient receptor potential channel" HUMAN MOLECULAR GENETICS, vol. 9, no. 17, 12 October 2000 (2000-10-12), pages 2471-2478, XP002176006 gene MLIV abstract page 2474, column 2, line 22 -page 2476, column 1, line 27 page 2477, column 1, line 30 - line 46	1,3,7,13
P,X	-& DATABASE EMBL, HEIDELBERG, FRG 'Online! 1 March 2001 (2001-03-01) SUN, M. ET AL.: "CDNA: FLJ22449 FIS, CLONE HRC09609 (MUCOLIPIN) (MUCOLIPIDOSIS TYPE IV PROTEIN) (MUCOLIPIN 1)" retrieved from HOMO SAPIENS Database accession no. Q9GZU1 XP002176010 the whole document	1,3,7

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Internatio pplication No
PCT/US 01/11442

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0017222 A	30-03-2000	AU 5928899 A	10-04-2000
		EP 1115735 A	18-07-2001
WO 0112662 A	22-02-2001	AU 6906800 A	13-03-2001

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 October 2001 (18.10.2001)

PCT

(10) International Publication Number
WO 01/077331 A1

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, 16/18, C12Q 1/68, G01N 33/53, A61K 38/17, 39/395, 48/00

(21) International Application Number: PCT/US01/11442

(22) International Filing Date: 6 April 2001 (06.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/544,797 7 April 2000 (07.04.2000) US

(71) Applicant *(for all designated States except US)*: **MILLENNIUM PHARMACEUTICALS, INC.** [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CURTIS, Rory, A., J.** [GB/US]; 31 Constitution Drive, Southborough, MA 01772 (US). **SILOS-SANTIAGO, Inmaculada** [ES/US]; 18 Hilliard Street, Cambridge, MA 02138 (US).

(74) Agents: **MANDRAGOURAS, Amy, E.**; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: HUMAN TRP-LIKE CALCIUM CHANNEL PROTEIN-2 (TLCC-2)

TRANSMEMBRANE SEGMENTS PREDICTED BY MEMSAT

START	END	ORIENT	SCORE
70	86	ins-->out	2.0
299	317	out-->ins	3.7
354	371	ins-->out	2.4
385	402	out-->ins	3.1
428	447	ins-->out	4.7
497	521	out-->ins	6.4

(57) Abstract: The invention provides isolated nucleic acids molecules, designated TLCC-2 nucleic acid molecules, which encode novel TRP-like calcium channel molecules. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing TLCC-2 nucleic acid molecules, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a TLCC-2 gene has been introduced or disrupted. The invention still further provides isolated TLCC-2 proteins, fusion proteins, antigenic peptides and anti-TLCC-2 antibodies. Diagnostic methods utilizing compositions of the invention are also provided.

>54420
MTAPAGPRGSETERLLTPNPGYGTQAGPSPAPPTPPEEEDLRRRLKYFFMSPCDKFRAKG
RKPCKLMLQVVKILVVTVQLILFGLSNQLAVTFREENTIAFRHLFLLGYSDGADDTFAAY
TREQLYQAIFHAVDQYLALPDVSLGRYAYVRGGGDPWTNGSGLALCQRYYHRGHVDPAND
TFDIDPMVVTDCIQVDPPERPPPPPSDDLTLLESSSSYKNLTLKFHKLVNVTIHFRLKTI
NLQSLINNEIPDCYTFSVLITFDNKAHSGRIPISLETQAHIQECKHPSVFQHGDNSFRLL
FDVVVILTCSLSFLLCARSLLRGFLLQNEFVGFMWRQRGRVISLWERLEFVNGWYILLVT
SDVLTISGTIMKIGIEAKNLASYDVCSILLGTSTLLVWVGVIRYLTFFHNYNILIATLRV
ALPSVMRFCCCVAVIYLGYCFCGWIVLGPYHVKFRSLSMVSECLFSLINGDDMFVTFAAM
QAQQGRSSLVWLFSQLYLYSFISLFIYMVLSLFIALITGAYDTIKHPGGAGAEESELQAY
IAQCQDSPTSGKFRRGSGSACSLLCCCGRDPSEEHSLLVN

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
with international search report

(48) **Date of publication of this corrected version:**
19 December 2002

(15) **Information about Correction:**
see PCT Gazette No. 51/2002 of 19 December 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

HUMAN TRP-LIKE CALCIUM CHANNEL PROTEIN-2 (TLCC-2)

Related Applications

The present application is a continuation-in-part of U.S. Patent Application Serial No. 09/544,797 entitled "54420, A NOVEL HUMAN CALCIUM CHANNEL", filed April 7, 2000. The content of the above-referenced patent application is incorporated herein by this reference in its entirety.

Background of the Invention

Calcium signaling has been implicated in the regulation of a variety of cellular responses, such as growth and differentiation. There are two general methods by which intracellular concentrations of calcium ions may be increased: calcium ions may be brought into the cell from the extracellular milieu through the use of specific channels in the cellular membrane, or calcium ions may be freed from intracellular stores, again being transported by specific membrane channels in the storage organelle. In the situation in which the intracellular stores of calcium have been depleted, a specific type of calcium channel, termed a 'capacitative calcium channel' or a 'store-operated calcium channel' (SOC), is activated in the plasma membrane to import calcium ions from the extracellular environment to the cytosol (for review, see Putney and McKay (1999) *BioEssays* 21:38-46).

Members of the capacitative calcium channel family include the calcium release-activated calcium current (CRAC) (Hoth and Penner (1992) *Nature* 355: 353-355), calcium release-activated nonselective cation current (CRANC) (Krause *et al.* (1996) *J. Biol. Chem.* 271: 32523-32528), and the transient receptor potential (TRP) proteins. There is no single electrophysological profile characteristic of the family; rather, a wide array of single channel conductances, cation selectivity, and current properties have been observed for different specific channels. Further, in several instances it has been demonstrated that homo- or heteropolymerization of the channel molecule may occur, further changing the channel properties from that of the single molecule. In general, though, these channels function similarly, in that they are calcium ion-permeable cation channels which become activated upon stimulation of phospholipase C_β by a G protein-coupled receptor. Depletion of intracellular calcium stores activate these channels by a

mechanism which is yet undefined, but which has been demonstrated to involve a diffusible factor using studies in which calcium stores were artificially depleted (*e.g.*, by the introduction of chelators into the cell, by activating phospholipase C_γ, or by inhibiting those enzymes responsible for pumping calcium ions into the stores or those enzymes responsible for maintaining resting intracellular calcium ion concentrations) (Putney, J.W., (1986) *Cell Calcium* 7: 1-12; Putney, J.W. (1990) *Cell Calcium* 11:611-624).

The TRP channel family is one of the best characterized members of the capacitative calcium channel group. These channels include transient receptor potential protein and homologues thereof (to date, seven homologs and splice variants have been identified in a variety of organisms), the vanilloid receptors (also known as the capsaicin receptors), stretch-inhibitable non-selective cation channel (SIC), olfactory, mechanosensitive channel, insulin-like growth factor I-regulated calcium channel, and vitamin D-responsive apical, epithelial calcium channel (ECaC), melastatin, and the polycystic kidney disease protein family (see, *e.g.*, Montell and Rubin (1989) *Neuron* 2:1313-1323; Caterina *et al.* (1997) *Nature* 389: 816-824; Suzuki *et al.* (1999) *J. Biol. Chem.* 274: 6330-6335; Kiselyov *et al.* (1998) *Nature* 396: 478-482; Hoenderop *et al.* (1999) *J. Biol. Chem.* 274: 8375-8378; and Chen et al. (1999) Nature 401(6751): 383-6). Each of these molecules is 700 or more amino acids in length (TRP and TRP homologs have 1300 or more amino acid residues), and shares certain conserved structural features. Predominant among these structural features are six transmembrane domains, with an additional hydrophobic loop present between the fifth and sixth transmembrane domains. It is believed that this loop is integral to the activity of the pore of the channel formed upon membrane insertion (Hardie and Minke (1993) *Trends Neurosci* 16: 371-376). TRP channel proteins also include one or more ankyrin domains and frequently display a proline-rich region at the N-terminus. Although found in disparate tissues and organisms, members of the TRP channel protein family all serve to transduce signals by means of calcium entry into cells, particularly pain signals (see, *e.g.*, McClesky and Gold (1999) *Annu. Rev. Physiol.* 61: 835-856), light signals (Hardie and Minke, *supra*), or olfactory signals (Colbert *et al.* (1997) *J. Neurosci* 17(21): 8259-8269). Thus, this family of molecules may play important roles in sensory signal transduction in general.

Vanilloid receptors (VRs) are non-selective cation channels that are structurally related to members of the TRP family of ion channels. These receptors have been proposed to mediate the entry of extracellular calcium into cells in response to the depletion of intracellular calcium stores. VRs are expressed in nociceptive neurons, as well as other cells types, and are activated by a variety of stimuli including noxious heat and protons. Capsaicin, which is a well-known agonist of VRs, induces pain behavior in humans and rodents. VR-1, a vanilloid receptor, was identified in rat sensory ganglia (Caterina M. J. et al., (1997) *Nature* 389:816-824). It has been shown that VR-1 knockout mice are impaired in their detection of painful heat, exhibit no vanilloid-evoked pain behavior, and show little thermal hypersensitivity after inflammation (Szallasi and Blumberg (1999) Pharmacol. Rev. 51:159-211; Tominaga, *et al.* (1998) *Neuron* 21:531; Caterina *et al.* (2000) *Science* 288:306).

Summary of the Invention

The present invention is based, at least in part, on the discovery of novel transient receptor potential (TRP) (*e.g.*, the calcium channel and/or vanilloid receptor) family members, referred to herein as TRP-like calcium channel or TLCC-2 nucleic acid and protein molecules. The TLCC-2 molecules of the present invention are useful as targets for developing modulating agents to regulate a variety of cellular processes, including membrane excitability, neurite outgrowth and synaptogenesis, signal transduction, cell proliferation, growth, differentiation, and migration, and nociception. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding TLCC-2 proteins or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of TLCC-2-encoding nucleic acids.

In one embodiment, a TLCC-2 nucleic acid molecule of the invention is at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the nucleotide sequence (*e.g.*, to the entire length of the nucleotide sequence) shown in SEQ ID NO:1 or 3 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a complement thereof.

In a preferred embodiment, the isolated nucleic acid molecule includes the nucleotide sequence shown SEQ ID NO:1 or 3, or a complement thereof. In another embodiment, the nucleic acid molecule includes SEQ ID NO:3 and nucleotides 1-140 of SEQ ID NO:1. In yet another embodiment, the nucleic acid molecule includes SEQ ID NO:3 and nucleotides 1884-2095 of SEQ ID NO:1. In another preferred embodiment, the nucleic acid molecule consists of the nucleotide sequence shown in SEQ ID NO:1 or 3. In another preferred embodiment, the nucleic acid molecule includes a fragment of at least 706 nucleotides (*e.g.*, 706 contiguous nucleotides) of the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, or a complement thereof.

In another embodiment, a TLCC-2 nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In a preferred embodiment, a TLCC-2 nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the entire length of the amino acid sequence of SEQ ID NO:2 or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

In another preferred embodiment, an isolated nucleic acid molecule encodes the amino acid sequence of human TLCC-2. In yet another preferred embodiment, the nucleic acid molecule includes a nucleotide sequence encoding a protein having the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, the nucleic acid molecule is at least 706 nucleotides in length. In a further preferred embodiment, the nucleic acid molecule is at least 706 nucleotides in length and encodes a protein having a TLCC-2 activity (as described herein).

Another embodiment of the invention features nucleic acid molecules, preferably TLCC-2 nucleic acid molecules, which specifically detect TLCC-2 nucleic acid molecules relative to nucleic acid molecules encoding non-TLCC-2 proteins. For example, in one embodiment, such a nucleic acid molecule is at least 706, 706-750, 750-800, 800-850, 850-900, 900-950, 950-1000, 1000-1050, 1050-1070, 1070-1100, 1100-

1150, 1150-1200, 1200-1250, 1250-1300, 1300-1350, 1350-1400, 1400-1450, 1450-1500, 1500-1550, 1550-1600, 1600-1650, 1650-1700, 1700-1750, 1750-1800, 1800-1850, 1850-1900, 1900-1950, 1950-2000, 2000-2050, 2050-2100 or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence shown in SEQ ID NO:1, the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number_____, or a complement thereof.

In preferred embodiments, the nucleic acid molecules are at least 15 (*e.g.*, 15 contiguous) nucleotides in length and hybridize under stringent conditions to nucleotides 1-27 of SEQ ID NO:1.

In other preferred embodiments, the nucleic acid molecule encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:1 or 3 under stringent conditions.

Another embodiment of the invention provides an isolated nucleic acid molecule which is antisense to a TLCC-2 nucleic acid molecule, *e.g.*, the coding strand of a TLCC-2 nucleic acid molecule.

Another aspect of the invention provides a vector comprising a TLCC-2 nucleic acid molecule. In certain embodiments, the vector is a recombinant expression vector. In another embodiment, the invention provides a host cell containing a vector of the invention. In yet another embodiment, the invention provides a host cell containing a nucleic acid molecule of the invention. The invention also provides a method for producing a protein, preferably a TLCC-2 protein, by culturing in a suitable medium, a host cell, *e.g.*, a mammalian host cell such as a non-human mammalian cell, of the invention containing a recombinant expression vector, such that the protein is produced.

Another aspect of this invention features isolated or recombinant TLCC-2 proteins and polypeptides. In one embodiment, an isolated TLCC-2 protein has one or more of the following domains: a transmembrane domain, a pore domain, and a proline rich domain. In a preferred embodiment, a TLCC-2 protein includes at least one or more of the following domains: a transmembrane domain, a pore domain, and a proline

rich domain, and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another preferred embodiment, a TLCC-2 protein includes at least one transmembrane domain and has a TLCC-2 activity (as described herein).

In yet another preferred embodiment, a TLCC-2 protein includes one or more of the following domains: a transmembrane domain, a pore domain, and a proline rich domain, and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3.

In another embodiment, the invention features fragments of the protein having the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 18 or more amino acids (*e.g.*, contiguous amino acids) of the amino acid sequence of SEQ ID NO:2, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number_____. In another embodiment, a TLCC-2 protein has the amino acid sequence of SEQ ID NO:2.

In another embodiment, the invention features a TLCC-2 protein which is encoded by a nucleic acid molecule consisting of a nucleotide sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to a nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof. This invention further features a TLCC-2 protein, which is encoded by a nucleic acid molecule consisting of a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof.

The proteins of the present invention or portions thereof, *e.g.*, biologically active portions thereof, can be operatively linked to a non-TLCC-2 polypeptide (*e.g.*, heterologous amino acid sequences) to form fusion proteins. The invention further features antibodies, such as monoclonal or polyclonal antibodies, that specifically bind proteins of the invention, preferably TLCC-2 proteins. In addition, the TLCC-2 proteins or biologically active portions thereof can be incorporated into pharmaceutical compositions, which optionally include pharmaceutically acceptable carriers.

In another aspect, the present invention provides a method for detecting the presence of a TLCC-2 nucleic acid molecule, protein, or polypeptide in a biological sample by contacting the biological sample with an agent capable of detecting a TLCC-2 nucleic acid molecule, protein, or polypeptide such that the presence of a TLCC-2 nucleic acid molecule, protein or polypeptide is detected in the biological sample.

In another aspect, the present invention provides a method for detecting the presence of TLCC-2 activity in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of TLCC-2 activity such that the presence of TLCC-2 activity is detected in the biological sample.

In another aspect, the invention provides a method for modulating TLCC-2 activity comprising contacting a cell capable of expressing TLCC-2 with an agent that modulates TLCC-2 activity such that TLCC-2 activity in the cell is modulated. In one embodiment, the agent inhibits TLCC-2 activity. In another embodiment, the agent stimulates TLCC-2 activity. In one embodiment, the agent is an antibody that specifically binds to a TLCC-2 protein. In another embodiment, the agent modulates expression of TLCC-2 by modulating transcription of a TLCC-2 gene or translation of a TLCC-2 mRNA. In yet another embodiment, the agent is a nucleic acid molecule having a nucleotide sequence that is antisense to the coding strand of a TLCC-2 mRNA or a TLCC-2 gene.

In one embodiment, the methods of the present invention are used to treat a subject having a disorder characterized by aberrant or unwanted TLCC-2 protein or nucleic acid expression or activity, *e.g.*, a pain disorder, by administering an agent which is a TLCC-2 modulator to the subject. In one embodiment, the TLCC-2 modulator is a TLCC-2 protein. In another embodiment the TLCC-2 modulator is a TLCC-2 nucleic acid molecule. In yet another embodiment, the TLCC-2 modulator is an antibody, ribozyme, peptide, peptidomimetic, antisense oligonucleotide, or other small molecule. In a preferred embodiment, the disorder characterized by aberrant or unwanted TLCC-2 protein or nucleic acid expression is a CNS disorder, such as a neurodegenerative disorder, *e.g.*, Alzheimer's disease, dementias related to Alzheimer's disease (such as Pick's disease), Parkinson's and other Lewy diffuse body diseases, multiple sclerosis, amyotrophic lateral sclerosis, progressive supranuclear palsy, epilepsy, Jakob-Creutzfieldt disease, or AIDS related dementia, familial infantile

convulsions, paroxysmal choreoathetosis; a disorder of the conveyance of sensory impulses from the periphery to the brain and/or conductance of motor impulses from the brain to the periphery; a psychiatric disorder (*e.g.*, depression, schizophrenic disorders, korsakoff's psychosis, mania, anxiety disorders, or phobic disorders; or a learning or memory disorder (*e.g.*, amnesia or age-related memory loss; or is a neurological disorder (*e.g.*, migraine).

In another embodiment, the disorder characterized by aberrant or unwanted TLCC-2 activity is a pain disorder, or a disorder characterized by misregulated pain signaling mechanisms.

In another embodiment, the disorder characterized by aberrant or unwanted TLCC-2 activity is a cell proliferation, growth, differentiation, or migration disorder.

The present invention also provides diagnostic assays for identifying the presence or absence of a genetic alteration characterized by at least one of (i) aberrant modification or mutation of a gene encoding a TLCC-2 protein; (ii) mis-regulation of the gene; and (iii) aberrant post-translational modification of a TLCC-2 protein, wherein a wild-type form of the gene encodes a protein with a TLCC-2 activity.

In another aspect the invention provides methods for identifying a compound that binds to or modulates the activity of a TLCC-2 protein, by providing an indicator composition comprising a TLCC-2 protein having TLCC-2 activity, contacting the indicator composition with a test compound, and determining the effect of the test compound on TLCC-2 activity in the indicator composition to identify a compound that modulates the activity of a TLCC-2 protein.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Drawings

Figure 1A-B depicts the cDNA sequence and predicted amino acid sequence of human TLCC-2. The nucleotide sequence corresponds to nucleic acids 1 to 2095 of SEQ ID NO:1. The amino acid sequence corresponds to amino acids 1 to 580 of SEQ ID NO:2. The coding region without the 3' untranslated region of the human TLCC-2 gene is shown in SEQ ID NO:3.

Figure 2 depicts a structural, hydrophobicity, and antigenicity analysis of the human TLCC-2 protein.

Figure 3 depicts the results of a search which was performed against the MEMSAT database and which resulted in the identification of six "transmembrane domains" in the human TLCC-2 protein (SEQ ID NO:2).

Figure 4 depicts the results of a search which was performed against the HMM database and which resulted in the identification of a "Fibronectin type III domain" in the human TLCC-2 protein (SEQ ID NO:2).

Figure 5 depicts the results of a search which was performed against the ProDom database and which resulted in the identification of a "31K RNA-4 protein domain" in the human TLCC-2 protein (SEQ ID NO:2).

Figure 6 depicts an alignment of the human TLCC-2 amino acid sequence (SEQ ID NO:2) with the amino acid sequences of human vanilloid receptor 1 and human vanilloid receptor 2 (Accession Numbers AAD26363 and AAD41724, set forth as SEQ ID NO:4 and SEQ ID NO:5, respectively), using the CLUSTAL W(1.74) alignment program.

Figure 7A-B depicts an alignment of the human TLCC-2 amino acid sequence (SEQ ID NO:2) with the amino acid sequence of polycystic kidney disease protein 2 from *Mus musculus* (Accession Number NP_032887, set forth as SEQ ID NO:6), using the CLUSTAL W(1.74) alignment program.

Figure 8A-B depicts an alignment of the human TLCC-2 amino acid sequence (SEQ ID NO:2) with the amino acid sequence of human melastatin (Accession Number AAC8000, set forth as SEQ ID NO:7), using the CLUSTAL W(1.74) alignment program.

Figure 9 depicts the results of a search performed against the Prosite database and which resulted in the identification of four N-glycosylation sites in the amino acid sequence of human TLCC-2 (SEQ ID NO:2).

Figure 10 is a graphic depiction of the relative levels of the human TLCC-2 mRNA expression in a human normal tissue panel, as determined using Taqman™ analysis.

Figure 11 is a graphic depiction of the relative levels of the human TLCC-2 mRNA expression in a human normal tissue panel, as determined using Taqman™ analysis.

Detailed Description of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as " TRP-like calcium channel" or "TLCC-2" nucleic acid and protein molecules, which are novel members of the ion channel, *e.g.,* calcium channel and/or vanilloid receptor, family. These novel molecules are capable of, for example, modulating an ion-channel mediated activity (*e.g.*, a calcium channel- and/or vanilloid receptor-mediated activity) in a cell, *e.g.*, a neuronal, skin, muscle (*e.g.*, cardiac muscle), or liver cell.

The present invention is based also in part, on the discovery that the novel TLCC-2 molecules of the present invention are expressed in the brain at high levels, and are also expressed in the skin, the spinal cord and dorsal root ganglia (DRG). Moreover, the novel TLCC-2 molecules of the invention are upregulated in animal models of pain. The TLCC-2 molecules of the invention are involved in nociception (*e.g.*, chemical, mechanical, or thermal nociception) and thereby modulate pain elicitation. Accordingly, the TLCC-2 molecules of the present invention act as targets for developing novel diagnostic targets and therapeutic agents to control pain and pain disorders.

As used herein, an "ion channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting signals in an electrically excitable cell, *e.g.*, a neuronal or muscle cell. Ion channels include vanilloid receptors, calcium channels, potassium channels, and sodium channels. As used herein, a "calcium channel" includes a protein or polypeptide which is involved in receiving, conducting, and trasmitting calcium ion-based signals in an electrically excitable cell. Calcium channels are calcium ion selective, and can determine membrane excitability (the ability of, for example, a neuronal cell to respond to a stimulus and to convert it into a sensory impulse). Calcium channels can also influence the resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation. Calcium channels are typically expressed in electrically excitable cells, *e.g.*, neuronal cells, and may form heteromultimeric structures (*e.g.*, composed of more than one type of

subunit). Calcium channels may also be found in non-excitable cells (*e.g.*, adipose cells or liver cells), where they may play a role in, *e.g.*, signal transduction.. Examples of calcium channels include the low-voltage-gated channels and the high-voltage-gated channels. Calcium channels are described in, for example, Davila *et al.* (1999) *Annals New York Academy of Sciences* 868:102-17 and McEnery, M.W. *et al.* (1998) *J. Bioenergetics and Biomembranes* 30(4): 409-418, the contents of which are incorporated herein by reference.

As used herein, a "vanilloid receptor" includes a non- selective cation channel that is structurally related to the TRP family of ion channels. Vanilloid receptors are also known as capsaisin receptors. Vanilloid receptors share several physical characteristics including an N-terminal cytoplasmic domain which contains three ankyrin repeats, six transmembrane domains, a pore-loop region located between transmembrane domains 5 and 6, and several kinase consensus sequences. Members of the vanilloid receptor (VR) family have been proposed to mediate the entry of extracellular calcium into cells, *e.g.*, in response to the depletion of intracellular calcium stores. VRs are typically expressed in nociceptive neurons among other cells types and are directly activated by harmful heat, extracellular protons, and vanilloid compounds. VRs may also be expressed in nonsensory tissues and may mediate inflammatory rather than acute thermal pain. Vanilloid receptors are described in, for example, Caterina, M.J. (1997) *Nature* 389:816-824 and Caterina, M.J. (2000) *Science* 288:306-313) the contents of which are incorporated herein by reference. As the TLCC-2 molecules of the present invention may modulate ion channel mediated activities (*e.g.*, calcium channel- and/or vanilloid receptor- mediated activities), they may be useful for developing novel diagnostic and therapeutic agents for ion channel associated disorders (*e.g.*, calcium channel and/or vanilloid receptor associated disorders).

As used herein, an "ion channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of ion channel (*e.g.*, calcium channel) and/or vanilloid receptor) mediated activity. For example, a "calcium channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of calcium channel mediated activity. A "vanilloid receptor associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of vanilloid receptor mediated activity. Ion channel associated disorders,

e.g., calcium channel associated disorders and/or vanilloid receptor associated disorders, include CNS disorders, such as cognitive and neurodegenerative disorders, examples of which include, but are not limited to, Alzheimer's disease, dementias related to Alzheimer's disease (such as Pick's disease), Parkinson's and other Lewy diffuse body diseases, senile dementia, Huntington's disease, Gilles de la Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis, progressive supranuclear palsy, epilepsy, Jakob-Creutzfieldt disease, or AIDS related dementia; autonomic function disorders such as hypertension and sleep disorders, and neuropsychiatric disorders, such as depression , schizophrenia, schizoaffective disorder, korsakoff's psychosis, mania, anxiety disorders, or phobic disorders; leaning or memory disorders, *e.g.*, amnesia or age-related memory loss, attention deficit disorder, psychoactive substance use disorders, anxiety, phobias, panic disorder, as well as bipolar affective disorder, *e.g.*, severe bipolar affective (mood) disorder (BP-1), and bipolar affective neurological disorders, *e.g.*, migraine and obesity. Further CNS-related disorders include, for example, those listed in the American Psychiatric Association's Diagnostic and Statistical manual of Mental Disorders (DSM), the most current version of which is incorporated herein by reference in its entirety.

Ion channel associated disorders, *e.g.*, calcium channel disorders and/or vanilloid receptor associated disorders, also include pain disorders. As used herein, the term "pain disorder" includes a disorder affecting pain signaling mechanisms. Pain disorders include disorders characterized by aberrant (*e.g.*, excessive or amplified) pain. Examples of pain disorders include posttherapeutic neuralgia, diabetic neuropathy, postmastectomy pain syndrome, stump pain, reflex sympathetic dystrophy, trigeminal neuralgia, neuropathic pain, orofacial neuropathic pain, osteoarthritis, rheumatoid arthritis, fibromyalgia syndrome, tension myalgia, Guillian-Barre syndrome, Meralgia paraesthetica, burning mouth syndrome, fibrocitis, myofascial pain syndrome, idiopathic pain disorder, temporomandibular joint syndrome, atypical odontalgia, loin pain, haematuria syndrome, non-cardiac chest pain, low back pain, chronic nonspecific pain, psychogenic pain, musculoskeletal pain disorder, chronic pelvic pain, nonorganic chronic headache, tension-type headache, cluster headache, migraine, complex regional pain syndrome, vaginismus, nerve trunk pain, somatoform pain disorder, cyclical mastalgia, chronic fatigue syndrome, multiple somatization syndrome, chronic pain

disorder, somatization disorder, Syndrome X, facial pain, idiopathic pain disorder, posttraumatic rheumatic pain modulation disorder (fibrositis syndrome), hyperalgesia, and Tangier disease.

As used herein, the term "pain signaling mechanisms" includes the cellular mechanisms involved in the development and regulation of pain, *e.g.*, pain elicited by noxious chemical, mechanical, or thermal stimuli, in a subject, *e.g.*, a mammal such as a human. In mammals, the initial detection of noxious chemical, mechanical, or thermal stimuli, a process referred to as "nociception", occurs predominantly at the peripheral terminals of specialized, small diameter sensory neurons. These sensory neurons transmit the information to the central nervous system, evoking a perception of pain or discomfort and initiating appropriate protective reflexes. The TLCC-2 molecules of the present invention may be present on these sensory neurons and, thus, may be involved in detecting these noxious chemical, mechanical, or thermal stimuli and transducing this information into membrane depolarization events. Thus, the TLCC-2 molecules, by participating in pain signaling mechanisms, may modulate pain elicitation and act as targets for developing novel diagnostic targets and therapeutic agents to control pain.

Ion channel associated disorders, *e.g.*, calcium channel disorders and/or vanilloid receptor associated disorders, also include cellular proliferation, growth, differentiation, or migration disorders. Cellular proliferation, growth, differentiation, or migration disorders include those disorders that affect cell proliferation, growth, differentiation, or migration processes. As used herein, a "cellular proliferation, growth, differentiation, or migration process" is a process by which a cell increases in number, size or content, by which a cell develops a specialized set of characteristics which differ from that of other cells, or by which a cell moves closer to or further from a particular location or stimulus. The TLCC-2 molecules of the present invention are involved in signal transduction mechanisms, which are known to be involved in cellular growth, differentiation, and migration processes. Thus, the TLCC-2 molecules may modulate cellular growth, differentiation, or migration, and may play a role in disorders characterized by aberrantly regulated growth, differentiation, or migration. Such disorders include cancer, *e.g.*, carcinoma, sarcoma, or leukemia; tumor angiogenesis and metastasis; skeletal dysplasia; neuronal deficiencies resulting from impaired neural induction and

patterning; hepatic disorders; cardiovascular disorders; and hematopoietic and/or myeloproliferative disorders.

As used herein, an "ion channel mediated activity" includes an activity which involves an ion channel, *e.g.*, an ion channel in a neuronal cell, a skin cell, a muscular cell, or a liver cell, associated with receiving, conducting, and transmitting signals, in, for example, the nervous system. Ion channel mediated activities (*e.g.*, calcium channel and/or vanilloid receptor mediated activities) include release of neurotransmitters or second messenger molecules (*e.g.*, dopamine or norepinephrine), from cells, *e.g.*, neuronal cells; mediation of entry of extracellular calcium into cells, *e.g.*, neuronal cells; modulation of resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation; participation in signal transduction pathways, and modulation of processes such as integration of sub-threshold synaptic responses and the conductance of back-propagating action potentials in, for example, neuronal cells (*e.g.*, changes in those action potentials resulting in a morphological or differentiative response in the cell).

The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. Such family members can be naturally or non-naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of human origin or alternatively, can contain homologues of non-human origin, *e.g.*, monkey proteins. Members of a family may also have common functional characteristics.

For example, the family of TLCC-2 proteins comprise at least one "transmembrane domain" and preferably six transmembrane domains. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 20-45 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a

transmembrane domain are hydrophobic, *e.g.*, leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W.N. *et al,* (1996) *Annual Rev. Neurosci.* 19: 235-263, the contents of which are incorporated herein by reference. Amino acid residues 70-86, 299-317, 354-371, 385-416, 428-447, and 497-521 of the TLCC-2 protein comprise transmembrane domains (see Figures 2 and 3). Accordingly, TLCC-2 proteins having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a transmembrane domain of human TLCC-2 are within the scope of the invention.

In another embodiment, a TLCC-2 molecule of the present invention is identified based on the presence of at least one pore domain between the fifth and sixth transmembrane domains. As used herein, the term "pore domain" includes an overall hydrophobic amino acid sequence which is located between two transmembrane domains of a calcium channel protein, preferably transmembrane domains 5 and 6, and which is believed to be a major determinant of ion selectivity and channel activity in calcium channels. Pore domains are described, for example in Vannier *et al.* (1998) *J. Biol. Chem.* 273: 8675-8679 and Phillips, A. M. *et al.* (1992) *Neuron* 8, 631-642, the contents of which are incorporated herein by reference. TLCC-2 molecules having at least one pore domain are within the scope of the invention. A pore domain may be found in the human TLCC-2 sequence (SEQ ID NO:2) at about residues 459-470 (Figure 2).

In another embodiment, a TLCC-2 molecule of the present invention is identified based on the presence of at least one N-glycosylation site. As used herein, the term "N-glycosylation site" includes an amino acid sequence of about 4 amino acid residues in length which serves as a glycosylation site. More preferably, an N-glycosylation site has the consensus sequence Asn-Xaa-Ser/Thr (where Xaa may be any amino acid) (SEQ ID NO:4). N-glycosylation sites are described in, for example, Prosite PDOC00001 (http://www.expasy.ch/cgi-bin/get-prodoc-entry?PDOC00001), the contents of which are incorporated herein by reference. Amino acid residues 159-162, 179-182, 220-223, and 230-233 of the TLCC-2 protein comprise N-glycosylation sites (see Figure 9). Accordingly, TLCC-2 proteins having at least one N-glycosylation site are within the scope of the invention.

In another embodiment, a TLCC-2 molecules of the present invention is identified based on the presence of a "proline rich domain" in the protein or corresponding nucleic acid molecule. As used herein, the term "proline rich domain" includes an amino acid sequence of about 4-6 amino acid residues in length having the general sequence X-Pro-X-X-Pro-X (where X can be any amino acid). Proline rich domains are usually located in a helical structure and bind through hydrophobic interactions to SH3 domains. SH3 domains recognize proline rich domains in both forward and reverse orientations. Proline rich domains are described in, for example, Sattler, M. *et al.* (1998) *Leukemia* 12: 637-644, the contents of which are incorporated herein by reference. Residues 1-37 of the amino acid sequence of human TLCC-2 (SEQ ID NO:2) contain proline-rich domains.

In a preferred embodiment, the TLCC-2 molecules of the invention include at least one transmembrane domain, at least one N-glycosylation site, at least one pore domain, and at least one proline rich domain.

Isolated proteins of the present invention, preferably TLCC-2 proteins, have an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2 or are encoded by a nucleotide sequence sufficiently identical to SEQ ID NO:1 or 3. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (*e.g.*, an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains have at least 30%, 40%, or 50% homology, preferably 60% homology, more preferably 70%-80%, and even more preferably 90-95% homology across the amino acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently identical. Furthermore, amino acid or nucleotide sequences which share at least 30%, 40%, or 50%, preferably 60%, more preferably 70-80%, or 90-95% homology and share a common functional activity are defined herein as sufficiently identical.

As used interchangeably herein, an "TLCC-2 activity", "biological activity of TLCC-2 " or "functional activity of TLCC-2", refers to an activity exerted by a TLCC-2 protein, polypeptide or nucleic acid molecule on a TLCC-2 responsive cell or tissue, or on a TLCC-2 protein substrate, as determined *in vivo*, or *in vitro*, according to standard techniques. In one embodiment, a TLCC-2 activity is a direct activity, such as an association with a TLCC-2-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with which a TLCC-2 protein binds or interacts in nature, such that TLCC-2-mediated function is achieved. A TLCC-2 target molecule can be a non-TLCC-2 molecule or a TLCC-2 protein or polypeptide of the present invention. In an exemplary embodiment, a TLCC-2 target molecule is a TLCC-2 ligand, *e.g.*, a calcium channel ligand. Alternatively, a TLCC-2 activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the TLCC-2 protein with a TLCC-2 ligand. The biological activities of TLCC-2 are described herein. For example, the TLCC-2 proteins of the present invention can have one or more of the following activities: (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials, (4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction, and (7) participate in nociception.

Accordingly, another embodiment of the invention features isolated TLCC-2 proteins and polypeptides having a TLCC-2 activity. Preferred proteins are TLCC-2 proteins having at least one or more of the following domains: a transmembrane domain, an N-glycosylation site, a pore domain, and a proline rich domain, and, preferably, a TLCC-2 activity.

Additional preferred proteins have one or more of the following domains: a transmembrane domain, an N-glycosylation site, a pore domain, and a proline rich domain, and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3.

The nucleotide sequence of the isolated human TLCC-2 cDNA and the predicted amino acid sequence of the human TLCC-2 polypeptide are shown in Figure 1 and in SEQ ID NOs:1 and 2, respectively. A plasmid containing the nucleotide sequence encoding human TLCC-2 was deposited with the American Type Culture Collection

(ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on _____ and assigned Accession Number _____. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The human TLCC-2 gene, which is approximately 2095 nucleotides in length, encodes a protein having a molecular weight of approximately 65.7 kD and which is approximately 580 amino acid residues in length.

Various aspects of the invention are described in further detail in the following subsections:

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode TLCC-2 proteins or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify TLCC-2-encoding nucleic acid molecules (*e.g.*, TLCC-2 mRNA) and fragments for use as PCR primers for the amplification or mutation of TLCC-2 nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

The term "isolated nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated TLCC-2 nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA

of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, as a hybridization probe, TLCC-2 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to TLCC-2 nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO:1. The sequence of SEQ ID NO:1 corresponds to the human TLCC-2 cDNA. This cDNA comprises sequences encoding the human TLCC-2 protein (*i.e.*, "the coding region", from nucleotides 141-1883), as well as 5' untranslated sequences (nucleotides 1-140) and 3' untranslated

sequences (nucleotides 1884-2095). Alternatively, the nucleic acid molecule can comprise only the coding region of SEQ ID NO:1 (*e.g.*, nucleotides 141-1883, corresponding to SEQ ID NO:3).

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ , or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the entire length of the nucleotide sequence shown in SEQ ID NO:1 or 3, or the entire length of the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a TLCC-2 protein, *e.g.*, a biologically active portion of a TLCC-2 protein. The nucleotide sequence determined from the cloning of the TLCC-2 gene allows for the generation of probes and primers designed for use in identifying and/or cloning other TLCC-2 family members, as well as TLCC-2 homologues from other species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes

under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, 75, 80, 85, 90, 95, or 100 or more consecutive nucleotides of a sense sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, of an anti-sense sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In one embodiment, a nucleic acid molecule of the present invention comprises a nucleotide sequence which is greater than 706, 706-750, 750-800, 800-850, 850-900, 900-950, 950-1000, 1000-1050, 1050-1070, 1070-1100, 1100-1150, 1150-1200, 1200-1250, 1250-1300, 1300-1350, 1350-1400, 1400-1450, 1450-1500, 1500-1550, 1550-1600, 1600-1650, 1650-1700, 1700-1750, 1750-1800, 1800-1850, 1850-1900, 1900-1950, 1950-2000, 2000-2050, 2050-2100 or more nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

Probes based on the TLCC-2 nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a TLCC-2 protein, such as by measuring a level of a TLCC-2-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting TLCC-2 mRNA levels or determining whether a genomic TLCC-2 gene has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion of a TLCC-2 protein" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, which encodes a polypeptide having a TLCC-2 biological activity (the biological activities of the TLCC-2 proteins are described herein), expressing the encoded portion of the TLCC-2 protein (*e.g.*, by recombinant

expression *in vitro*) and assessing the activity of the encoded portion of the TLCC-2 protein.

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, due to degeneracy of the genetic code and thus encode the same TLCC-2 proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:2.

In addition to the TLCC-2 nucleotide sequences shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the TLCC-2 proteins may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the TLCC-2 genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a TLCC-2 protein, preferably a mammalian TLCC-2 protein, and can further include non-coding regulatory sequences, and introns.

Allelic variants of human TLCC-2 include both functional and non-functional TLCC-2 proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the human TLCC-2 protein that maintain the ability to bind a TLCC-2 ligand or substrate and/or modulate membrane excitability or signal transduction. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2, or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human TLCC-2 protein that do not have the ability to form functional calcium channels or to modulate membrane excitability. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion or

premature truncation of the amino acid sequence of SEQ ID NO:2, or a substitution, insertion or deletion in critical residues or critical regions.

The present invention further provides non-human orthologues of the human TLCC-2 proteins. Orthologues of the human TLCC-2 protein are proteins that are isolated from non- non-human organisms and possess the same TLCC-2 ligand binding and/or modulation of membrane excitation mechanisms of the human TLCC-2 protein. Orthologues of the human TLCC-2 protein can readily be identified as comprising an amino acid sequence that is substantially identical to SEQ ID NO:2.

Moreover, nucleic acid molecules encoding other TLCC-2 family members and, thus, which have a nucleotide sequence which differs from the TLCC-2 sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, another TLCC-2 cDNA can be identified based on the nucleotide sequence of human TLCC-2. Moreover, nucleic acid molecules encoding TLCC-2 proteins from different species, and which, thus, have a nucleotide sequence which differs from the TLCC-2 sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, a mouse TLCC-2 cDNA can be identified based on the nucleotide sequence of a human TLCC-2.

Nucleic acid molecules corresponding to natural allelic variants and homologues of the TLCC-2 cDNAs of the invention can be isolated based on their homology to the TLCC-2 nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Nucleic acid molecules corresponding to natural allelic variants and homologues of the TLCC-2 cDNAs of the invention can further be isolated by mapping to the same chromosome or locus as the TLCC-2 gene.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15, 20, 25, 30 or more nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In other embodiment, the nucleic acid is at least 706, 706-750, 750-800, 800-850, 850-900, 900-950, 950-1000, 1000-

1050, 1050-1070, 1070-1100, 1100-1150, 1150-1200, 1200-1250, 1250-1300, 1300-1350, 1350-1400, 1400-1450, 1450-1500, 1500-1550, 1550-1600, 1600-1650, 1650-1700, 1700-1750, 1750-1800, 1800-1850, 1850-1900, 1900-1950, 1950-2000, 2000-2050, 2050-2100 or more nucleotides in length.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45° C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length,

$T_m(°C) = 2(\#\text{ of A + T bases}) + 4(\#\text{ of G + C bases})$. For hybrids between 18 and 49 base pairs in length, $T_m(°C) = 81.5 + 16.6(\log_{10}[Na^+]) + 0.41(\%G+C) - (600/N)$, where N is the number of bases in the hybrid, and $[Na^+]$ is the concentration of sodium ions in the hybridization buffer ($[Na^+]$ for 1xSSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS), chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH_2PO_4, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH_2PO_4, 1% SDS at 65°C, see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).

Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1 or 3 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In addition to naturally-occurring allelic variants of the TLCC-2 sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby leading to changes in the amino acid sequence of the encoded TLCC-2 proteins, without altering the functional ability of the TLCC-2 proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of TLCC-2 (*e.g.*, the sequence of SEQ ID NO:2) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the TLCC-2 proteins of the present invention, *e.g.*, those present in a

transmembrane domain, are predicted to be particularly unamenable to alteration. Furthermore, additional amino acid residues that are conserved between the TLCC-2 proteins of the present invention and other members of the TLCC-2 family are not likely to be amenable to alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding TLCC-2 proteins that contain changes in amino acid residues that are not essential for activity. Such TLCC-2 proteins differ in amino acid sequence from SEQ ID NO:2, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:2.

An isolated nucleic acid molecule encoding a TLCC-2 protein identical to the protein of SEQ ID NO:2, can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a TLCC-2 protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be

introduced randomly along all or part of a TLCC-2 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for TLCC-2 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In a preferred embodiment, a mutant TLCC-2 protein can be assayed for the ability to: (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials, (4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction,(7) participate in nociception, and (8) modulate pain signaling mechanisms.

In addition to the nucleic acid molecules encoding TLCC-2 proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire TLCC-2 coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding TLCC-2. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the coding region of human TLCC-2 corresponds to SEQ ID NO:3). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding TLCC-2. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding TLCC-2 disclosed herein (*e.g.*, SEQ ID NO:3), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of TLCC-2 mRNA, but more preferably is an

oligonucleotide which is antisense to only a portion of the coding or noncoding region of TLCC-2 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of TLCC-2 mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a TLCC-2 protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by

conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids. Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave TLCC-2 mRNA transcripts to thereby inhibit translation of TLCC-2 mRNA. A ribozyme having specificity for a TLCC-2-encoding nucleic acid can be designed based upon the nucleotide sequence of a TLCC-2 cDNA disclosed herein (*i.e.*, SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a TLCC-2-

encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, TLCC-2 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, TLCC-2 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the TLCC-2 (*e.g.*, the TLCC-2 promoter and/or enhancers; *e.g.*, nucleotides 1-137 of SEQ ID NO:1) to form triple helical structures that prevent transcription of the TLCC-2 gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

In yet another embodiment, the TLCC-2 nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.* *Proc. Natl. Acad. Sci.* 93: 14670-675.

PNAs of TLCC-2 nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of TLCC-2 nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (*e.g.*, by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.*, S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

In another embodiment, PNAs of TLCC-2 can be modified, (*e.g.*, to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of TLCC-2 nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (*e.g.*, RNAse H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, *e.g.*, Krol *et al.* (1988) *Bio-Techniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (*e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

Alternatively, the expression characteristics of an endogenous TLCC-2 gene within a cell line or microorganism may be modified by inserting a heterologous DNA regulatory element into the genome of a stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous TLCC-2 gene. For example, an endogenous TLCC-2 gene which is normally "transcriptionally silent", *i.e.*, a TLCC-2 gene which is normally not expressed, or is expressed only at very low levels in a cell line or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, a transcriptionally silent, endogenous TLCC-2 gene may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such that it is operatively linked with an endogenous TLCC-2 gene, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described, *e.g.*, in Chappel, U.S. Patent No. 5,272,071; PCT publication No. WO 91/06667, published May 16, 1991.

II. Isolated TLCC-2 Proteins and Anti-TLCC-2 Antibodies

One aspect of the invention pertains to isolated TLCC-2 proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-TLCC-2 antibodies. In one embodiment, native TLCC-2 proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, TLCC-2 proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a TLCC-2 protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the TLCC-2 protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of TLCC-2 protein in which the protein is separated from cellular components of the cells from which it is

isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of TLCC-2 protein having less than about 30% (by dry weight) of non-TLCC-2 protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-TLCC-2 protein, still more preferably less than about 10% of non-TLCC-2 protein, and most preferably less than about 5% non-TLCC-2 protein. When the TLCC-2 protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of TLCC-2 protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of TLCC-2 protein having less than about 30% (by dry weight) of chemical precursors or non-TLCC-2 chemicals, more preferably less than about 20% chemical precursors or non-TLCC-2 chemicals, still more preferably less than about 10% chemical precursors or non-TLCC-2 chemicals, and most preferably less than about 5% chemical precursors or non-TLCC-2 chemicals.

As used herein, a "biologically active portion" of a TLCC-2 protein includes a fragment of a TLCC-2 protein which participates in an interaction between a TLCC-2 molecule and a non-TLCC-2 molecule. Biologically active portions of a TLCC-2 protein include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the TLCC-2 protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2, which include less amino acids than the full length TLCC-2 proteins, and exhibit at least one activity of a TLCC-2 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the TLCC-2 protein, *e.g.*, modulating membrane excitation mechanisms. A biologically active portion of a TLCC-2 protein can be a polypeptide which is, for example, 25, 30, 35, 40, 45, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 274, 500, 525, 550, 575, 580, or more amino acids in length. Biologically active

portions of a TLCC-2 protein can be used as targets for developing agents which modulate a TLCC-2 mediated activity, *e.g.*, a membrane excitation mechanism.

In one embodiment, a biologically active portion of a TLCC-2 protein comprises at least one transmembrane domain. It is to be understood that a preferred biologically active portion of a TLCC-2 protein of the present invention comprises at least one or more of the following domains: a transmembrane domain, an N-glycosylation site, a pore domain, and a proline rich domain. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native TLCC-2 protein.

In a preferred embodiment, the TLCC-2 protein has an amino acid sequence shown in SEQ ID NO:2. In other embodiments, the TLCC-2 protein is substantially identical to SEQ ID NO:2, and retains the functional activity of the protein of SEQ ID NO:2, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the TLCC-2 protein is a protein which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:2.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (*e.g.*, when aligning a second sequence to the TLCC-2 amino acid sequence of SEQ ID NO:2 having 580 amino acid residues, at least 50, preferably at least 100, more preferably at least 200, even more preferably at least 300, and even more preferably at least 400 or 500 or more amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide

as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (*Comput. Appl. Biosci.*, 4:11-17 (1988)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to TLCC-2 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 100, wordlength = 3, and a Blosum62 matrix to obtain amino acid sequences homologous to TLCC-2 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.*

25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov.

The invention also provides TLCC-2 chimeric or fusion proteins. As used herein, a TLCC-2 "chimeric protein" or "fusion protein" comprises a TLCC-2 polypeptide operatively linked to a non-TLCC-2 polypeptide. A "TLCC-2 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to TLCC-2, whereas a "non-TLCC-2 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the TLCC-2 protein, *e.g.*, a protein which is different from the TLCC-2 protein and which is derived from the same or a different organism. Within a TLCC-2 fusion protein the TLCC-2 polypeptide can correspond to all or a portion of a TLCC-2 protein. In a preferred embodiment, a TLCC-2 fusion protein comprises at least one biologically active portion of a TLCC-2 protein. In another preferred embodiment, a TLCC-2 fusion protein comprises at least two biologically active portions of a TLCC-2 protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the TLCC-2 polypeptide and the non-TLCC-2 polypeptide are fused in-frame to each other. The non-TLCC-2 polypeptide can be fused to the N-terminus or C-terminus of the TLCC-2 polypeptide.

For example, in one embodiment, the fusion protein is a GST-TLCC-2 fusion protein in which the TLCC-2 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant TLCC-2.

In another embodiment, the fusion protein is a TLCC-2 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of TLCC-2 can be increased through the use of a heterologous signal sequence.

The TLCC-2 fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The TLCC-2 fusion proteins can be used to affect the bioavailability of a TLCC-2 substrate. Use of TLCC-2 fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a TLCC-2 protein;

(ii) mis-regulation of the TLCC-2 gene; and (iii) aberrant post-translational modification of a TLCC-2 protein.

Moreover, the TLCC-2-fusion proteins of the invention can be used as immunogens to produce anti-TLCC-2 antibodies in a subject, to purify TLCC-2 ligands and in screening assays to identify molecules which inhibit the interaction of TLCC-2 with a TLCC-2 substrate.

Preferably, a TLCC-2 chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A TLCC-2-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the TLCC-2 protein.

The present invention also pertains to variants of the TLCC-2 proteins which function as either TLCC-2 agonists (mimetics) or as TLCC-2 antagonists. Variants of the TLCC-2 proteins can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a TLCC-2 protein. An agonist of the TLCC-2 proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a TLCC-2 protein. An antagonist of a TLCC-2 protein can inhibit one or more of the activities of the naturally occurring form of the TLCC-2 protein by, for example, competitively modulating a TLCC-2-mediated activity of a TLCC-2 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the

biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the TLCC-2 protein.

In one embodiment, variants of a TLCC-2 protein which function as either TLCC-2 agonists (mimetics) or as TLCC-2 antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a TLCC-2 protein for TLCC-2 protein agonist or antagonist activity. In one embodiment, a variegated library of TLCC-2 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of TLCC-2 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential TLCC-2 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of TLCC-2 sequences therein. There are a variety of methods which can be used to produce libraries of potential TLCC-2 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential TLCC-2 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477.

In addition, libraries of fragments of a TLCC-2 protein coding sequence can be used to generate a variegated population of TLCC-2 fragments for screening and subsequent selection of variants of a TLCC-2 protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a TLCC-2 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which

encodes N-terminal, C-terminal and internal fragments of various sizes of the TLCC-2 protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of TLCC-2 proteins. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify TLCC-2 variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA 89*:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

In one embodiment, cell based assays can be exploited to analyze a variegated TLCC-2 library. For example, a library of expression vectors can be transfected into a cell line, *e.g.*, an endothelial cell line, which ordinarily responds to TLCC-2 in a particular TLCC-2 substrate-dependent manner. The transfected cells are then contacted with TLCC-2 and the effect of expression of the mutant on signaling by the TLCC-2 substrate can be detected, *e.g.*, by monitoring intracellular calcium, IP3, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, or the activity of a TLCC-2-regulated transcription factor. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the TLCC-2 substrate, and the individual clones further characterized.

An isolated TLCC-2 protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind TLCC-2 using standard techniques for polyclonal and monoclonal antibody preparation. A full-length TLCC-2 protein can be used or, alternatively, the invention provides antigenic peptide fragments of TLCC-2 for use as immunogens. The antigenic peptide of TLCC-2 comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2 and encompasses an

epitope of TLCC-2 such that an antibody raised against the peptide forms a specific immune complex with TLCC-2. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of TLCC-2 that are located on the surface of the protein, *e.g.*, hydrophilic regions, as well as regions with high antigenicity (see, for example, Figure 2).

A TLCC-2 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed TLCC-2 protein or a chemically synthesized TLCC-2 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic TLCC-2 preparation induces a polyclonal anti-TLCC-2 antibody response.

Accordingly, another aspect of the invention pertains to anti-TLCC-2 antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as TLCC-2. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind TLCC-2. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of TLCC-2. A monoclonal antibody composition thus typically displays a single binding affinity for a particular TLCC-2 protein with which it immunoreacts.

Polyclonal anti-TLCC-2 antibodies can be prepared as described above by immunizing a suitable subject with a TLCC-2 immunogen. The anti-TLCC-2 antibody titer in the immunized subject can be monitored over time by standard techniques, such

as with an enzyme linked immunosorbent assay (ELISA) using immobilized TLCC-2. If desired, the antibody molecules directed against TLCC-2 can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the anti-TLCC-2 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem* .255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally R. H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 54:387-402; M. L. Gefter *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a TLCC-2 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds TLCC-2.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-TLCC-2 monoclonal antibody (see, *e.g.*, G. Galfre *et al.* (1977) *Nature* 266:55052; Gefter *et al.* *Somatic Cell Genet.*, cited *supra*; Lerner, *Yale J. Biol. Med.*, cited *supra*; Kenneth, *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing

hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind TLCC-2, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-TLCC-2 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with TLCC-2 to thereby isolate immunoglobulin library members that bind TLCC-2. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92/15679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al. Nature* (1990) 348:552-554.

Additionally, recombinant anti-TLCC-2 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-TLCC-2 antibody (*e.g.*, monoclonal antibody) can be used to isolate TLCC-2 by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-TLCC-2 antibody can facilitate the purification of natural TLCC-2 from cells and of recombinantly produced TLCC-2 expressed in host cells. Moreover, an anti-TLCC-2 antibody can be used to detect TLCC-2 protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the TLCC-2 protein. Anti-TLCC-2 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, ß-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable

fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a TLCC-2 protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of

interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, TLCC-2 proteins, mutant forms of TLCC-2 proteins, fusion proteins, and the like).

The recombinant expression vectors of the invention can be designed for expression of TLCC-2 proteins in prokaryotic or eukaryotic cells. For example, TLCC-2 proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion

moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be utilized in TLCC-2 activity assays, (*e.g.*, direct assays or competitive assays described in detail below), or to generate antibodies specific for TLCC-2 proteins, for example. In a preferred embodiment, a TLCC-2 fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (*e.g.*, six (6) weeks).

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118).

Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the TLCC-2 expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

Alternatively, TLCC-2 proteins can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters

(*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to TLCC-2 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a TLCC-2 nucleic acid molecule of the invention is introduced, *e.g.*, a TLCC-2 nucleic acid molecule within a recombinant expression vector or a TLCC-2 nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or

environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a TLCC-2 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a TLCC-2 protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) a TLCC-2 protein. Accordingly, the invention further provides methods for producing a TLCC-2 protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a TLCC-2 protein has been introduced) in a suitable medium such that a TLCC-2 protein is produced. In

another embodiment, the method further comprises isolating a TLCC-2 protein from the medium or the host cell.

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which TLCC-2-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous TLCC-2 sequences have been introduced into their genome or homologous recombinant animals in which endogenous TLCC-2 sequences have been altered. Such animals are useful for studying the function and/or activity of a TLCC-2 and for identifying and/or evaluating modulators of TLCC-2 activity.. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous TLCC-2 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a TLCC-2-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The TLCC-2 cDNA sequence of SEQ ID NO:1 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human TLCC-2 gene, such as a mouse or rat TLCC-2 gene, can be used as a transgene. Alternatively, a TLCC-2 gene homologue, such as another TLCC-2 family member, can be isolated based on hybridization to the TLCC-2 cDNA sequences of SEQ ID NO:1 or 3, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____ (described further in subsection I above) and used as a

transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a TLCC-2 transgene to direct expression of a TLCC-2 protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a TLCC-2 transgene in its genome and/or expression of TLCC-2 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a TLCC-2 protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a TLCC-2 gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the TLCC-2 gene. The TLCC-2 gene can be a human gene (*e.g.*, the cDNA of SEQ ID NO:3), but more preferably, is a non-human homologue of a human TLCC-2 gene (*e.g.*, a cDNA isolated by stringent hybridization with the nucleotide sequence of SEQ ID NO:1). For example, a mouse TLCC-2 gene can be used to construct a homologous recombination nucleic acid molecule, *e.g.*, a vector, suitable for altering an endogenous TLCC-2 gene in the mouse genome. In a preferred embodiment, the homologous recombination nucleic acid molecule is designed such that, upon homologous recombination, the endogenous TLCC-2 gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the homologous recombination nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous TLCC-2 gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous TLCC-2 protein). In the homologous recombination nucleic acid molecule, the altered portion of the TLCC-2 gene is flanked at its 5' and 3'

ends by additional nucleic acid sequence of the TLCC-2 gene to allow for homologous recombination to occur between the exogenous TLCC-2 gene carried by the homologous recombination nucleic acid molecule and an endogenous TLCC-2 gene in a cell, *e.g.*, an embryonic stem cell. The additional flanking TLCC-2 nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the homologous recombination nucleic acid molecule (see, *e.g.*, Thomas, K.R. and Capecchi, M. R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid molecule is introduced into a cell, *e.g.*, an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced TLCC-2 gene has homologously recombined with the endogenous TLCC-2 gene are selected (see *e.g.*, Li, E. *et al.* (1992) *Cell* 69:915). The selected cells can then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see *e.g.*, Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, *e.g.*, vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and WO 93/04169 by Berns *et al.*

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a

selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0 phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

IV. Pharmaceutical Compositions

The TLCC-2 nucleic acid molecules, fragments of TLCC-2 proteins, and anti-TLCC-2 antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions

used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a fragment of a TLCC-2 protein or an anti-TLCC-2 antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the

use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While

compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

In a preferred example, a subject is treated with antibody, protein, or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody, protein, or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in

dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (*i.e*,. including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention.

Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (*e.g.*, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram.

It isfurthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the

activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

Further, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon *et al.*, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld *et al.* (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera *et al.* (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin *et al.* (eds.), pp. 303-16 (Academic Press 1985), and Thorpe *et al.*, "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982). Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells; *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (*e.g.*, therapeutic and

prophylactic). As described herein, a TLCC-2 protein of the invention has one or more of the following activities: (1) modulates membrane excitability, (2) influences the resting potential of membranes, (3) modulates wave forms and frequencies of action potentials, (4) modulates thresholds of excitation, (5) modulates neurite outgrowth and synaptogenesis, (6) modulates signal transduction, and (7) participates in nociception.

The isolated nucleic acid molecules of the invention can be used, for example, to express TLCC-2 protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect TLCC-2 mRNA (*e.g.*, in a biological sample) or a genetic alteration in a TLCC-2 gene, and to modulate TLCC-2 activity, as described further below. The TLCC-2 proteins can be used to treat disorders characterized by insufficient or excessive production of a TLCC-2 substrate or production of TLCC-2 inhibitors. In addition, the TLCC-2 proteins can be used to screen for naturally occurring TLCC-2 substrates, to screen for drugs or compounds which modulate TLCC-2 activity, as well as to treat disorders characterized by insufficient or excessive production of TLCC-2 protein or production of TLCC-2 protein forms which have decreased, aberrant or unwanted activity compared to TLCC-2 wild type protein (*e.g.*, CNS disorders (such as neurodegenerative disorders), pain disorders, or disorders of cellular growth, differentiation, or migration. Moreover, the anti-TLCC-2 antibodies of the invention can be used to detect and isolate TLCC-2 proteins, to regulate the bioavailability of TLCC-2 proteins, and modulate TLCC-2 activity.

A. Screening Assays:

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) which bind to TLCC-2 proteins, have a stimulatory or inhibitory effect on, for example, TLCC-2 expression or TLCC-2 activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of TLCC-2 substrate.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a TLCC-2 protein or polypeptide or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a

TLCC-2 protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a TLCC-2 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate TLCC-2 activity is determined. Determining the ability of the test compound to modulate TLCC-2 activity can be accomplished by monitoring, for example, intracellular calcium, IP3, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, or the activity of a TLCC-2-regulated transcription factor. The cell, for example, can be of mammalian origin, *e.g.*, a neuronal cell, or a liver cell.

The ability of the test compound to modulate TLCC-2 binding to a substrate or to bind to TLCC-2 can also be determined. Determining the ability of the test compound to modulate TLCC-2 binding to a substrate can be accomplished, for example, by coupling the TLCC-2 substrate with a radioisotope or enzymatic label such that binding of the TLCC-2 substrate to TLCC-2 can be determined by detecting the labeled TLCC-2 substrate in a complex. Alternatively, TLCC-2 could be coupled with a radioisotope or enzymatic label to monitor the ability of a test compound to modulate TLCC-2 binding to a TLCC-2 substrate in a complex. Determining the ability of the test compound to bind TLCC-2 can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to TLCC-2 can be determined by detecting the labeled TLCC-2 compound in a complex. For example, compounds (*e.g.*, TLCC-2 substrates) can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a compound (*e.g.*, a TLCC-2 substrate) to interact with TLCC-2 without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with TLCC-2 without the labeling of either the compound or the TLCC-2. McConnell, H. M. *et al.* (1992) *Science* 257:1906-1912. As used herein, a "microphysiometer" (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and TLCC-2.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a TLCC-2 target molecule (*e.g.*, a TLCC-2 substrate) with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the TLCC-2 target molecule. Determining the ability of the test compound to modulate the activity of a TLCC-2 target molecule can be

accomplished, for example, by determining the ability of the TLCC-2 protein to bind to or interact with the TLCC-2 target molecule.

Determining the ability of the TLCC-2 protein, or a biologically active fragment thereof, to bind to or interact with a TLCC-2 target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the TLCC-2 protein to bind to or interact with a TLCC-2 target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.*, intracellular Ca^{2+}, diacylglycerol, IP_3, and the like), detecting catalytic/enzymatic activity of the target using an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a target-regulated cellular response.

In yet another embodiment, an assay of the present invention is a cell-free assay in which a TLCC-2 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the TLCC-2 protein or biologically active portion thereof is determined. Preferred biologically active portions of the TLCC-2 proteins to be used in assays of the present invention include fragments which participate in interactions with non-TLCC-2 molecules, *e.g.*, fragments with high surface probability scores (see, for example, Figure 2). Binding of the test compound to the TLCC-2 protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the TLCC-2 protein or biologically active portion thereof with a known compound which binds TLCC-2 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TLCC-2 protein, wherein determining the ability of the test compound to interact with a TLCC-2 protein comprises determining the ability of the test compound to preferentially bind to TLCC-2 or biologically active portion thereof as compared to the known compound.

In another embodiment, the assay is a cell-free assay in which a TLCC-2 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the TLCC-2 protein or biologically active portion thereof is determined. Determining the ability of

the test compound to modulate the activity of a TLCC-2 protein can be accomplished, for example, by determining the ability of the TLCC-2 protein to bind to a TLCC-2 target molecule by one of the methods described above for determining direct binding. Determining the ability of the TLCC-2 protein to bind to a TLCC-2 target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In an alternative embodiment, determining the ability of the test compound to modulate the activity of a TLCC-2 protein can be accomplished by determining the ability of the TLCC-2 protein to further modulate the activity of a downstream effector of a TLCC-2 target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

In yet another embodiment, the cell-free assay involves contacting a TLCC-2 protein or biologically active portion thereof with a known compound which binds the TLCC-2 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the TLCC-2 protein, wherein determining the ability of the test compound to interact with the TLCC-2 protein comprises determining the ability of the TLCC-2 protein to preferentially bind to or modulate the activity of a TLCC-2 target molecule.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either TLCC-2 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a TLCC-2 protein, or interaction of a TLCC-2 protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion

protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/ TLCC-2 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or TLCC-2 protein, and the mixture incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of TLCC-2 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a TLCC-2 protein or a TLCC-2 target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated TLCC-2 protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with TLCC-2 protein or target molecules but which do not interfere with binding of the TLCC-2 protein to its target molecule can be derivatized to the wells of the plate, and unbound target or TLCC-2 protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the TLCC-2 protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the TLCC-2 protein or target molecule.

In another embodiment, modulators of TLCC-2 expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of TLCC-2 mRNA or protein in the cell is determined. The level of expression of TLCC-2 mRNA or protein in the presence of the candidate compound is compared to the level of expression of TLCC-2 mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of TLCC-2 expression

based on this comparison. For example, when expression of TLCC-2 mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of TLCC-2 mRNA or protein expression. Alternatively, when expression of TLCC-2 mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of TLCC-2 mRNA or protein expression. The level of TLCC-2 mRNA or protein expression in the cells can be determined by methods described herein for detecting TLCC-2 mRNA or protein.

In yet another aspect of the invention, the TLCC-2 proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with TLCC-2 ("TLCC-2-binding proteins" or "TLCC-2-bp") and are involved in TLCC-2 activity. Such TLCC-2-binding proteins are also likely to be involved in the propagation of signals by the TLCC-2 proteins or TLCC-2 targets as, for example, downstream elements of a TLCC-2-mediated signaling pathway. Alternatively, such TLCC-2-binding proteins are likely to be TLCC-2 inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a TLCC-2 protein is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a TLCC-2-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell

colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the TLCC-2 protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a TLCC-2 protein can be confirmed *in vivo*, *e.g.*, in an animal such as an animal model for cellular pain.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a TLCC-2 modulating agent, an antisense TLCC-2 nucleic acid molecule, a TLCC-2-specific antibody, or a TLCC-2-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

Models for studying pain *in vivo* include rat models of neuropathic pain caused by methods such as intraperitoneal administration of Taxol (Authier *et al.* (2000) *Brain Res.* 887:239-249), chronic constriction injury (CCI), partial sciatic nerve transection (Linenlaub and Sommer (2000) *Pain* 89:97-106), transection of the tibial and sural nerves (Lee *et al.* (2000) *Neurosci. Lett.* 291:29-32), the spared nerve injury model (Decosterd and Woolf (2000) *Pain* 87:149-158), cuffing the sciatic nerve (Pitcher and Henry (2000) *Eur. J. Neurosci.* 12:2006-2020), unilateral tight ligation (Esser and Sawynok (2000) *Eur. J. Pharmacol.* 399:131-139), L5 spinal nerve ligation (Honroe *et al.* (2000) *Neurosci.* 98:585-598), and photochemically induced ischemic nerve injury (Hao *et al.* (2000) *Exp. Neurol.* 163:231-238); rat models of nociceptive pain caused by methods such as the Chung Method, the Bennett Method, and intraperitoneal administration of complete Freund's adjuvant (CFA) (Abdi *et al.* (2000) *Anesth. Analg.* 91:955-959); rat models of post-incisional pain caused by incising the skin and fascia of a hind paw (Olivera and Prado (2000) *Braz. J. Med. Biol. Res.* 33:957-960); rat models of cancer pain caused by methods such as injecting osteolytic sarcoma cells into the

femur (Honroe *et al.* (2000) *Neurosci.* 98:585-598); and rat models of visceral pain caused by methods such as intraperitoneal administration of cyclophosphamide.

Various methods of determining an animal's response to pain are known in the art. Examples of such methods include, but are not limited to brief intense exposure to a focused heat source, administration of a noxious chemical subcutaneously, the tail flick test, the hot plate test, the formalin test, Von Frey threshold, and testing for stress-induced analgesia (*et al.*, by restraint, foot shock, and/or cold water swim) (Crawley (2000) *What's Wrong With My Mouse?* Wiley-Liss pp. 72-75).

B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

1. Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the TLCC-2 nucleotide sequences, described herein, can be used to map the location of the TLCC-2 genes on a chromosome. The mapping of the TLCC-2 sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, TLCC-2 genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the TLCC-2 nucleotide sequences. Computer analysis of the TLCC-2 sequences can be used to predict primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the TLCC-2 sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but human cells can, the one human chromosome that contains the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. (D'Eustachio P. *et al.* (1983) *Science* 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the TLCC-2 nucleotide sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a TLCC-2 sequence to its chromosome include *in situ* hybridization (described in Fan, Y. *et al.* (1990) *Proc. Natl. Acad. Sci. USA*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical such as colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this

technique, see Verma *et al.*, Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland, J. *et al.* (1987) *Nature*, 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the TLCC-2 gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

2. Tissue Typing

The TLCC-2 sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations

of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the TLCC-2 nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The TLCC-2 nucleotide sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from TLCC-2 nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

3. Use of TLCC-2 Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, *e.g.*, hair or skin, or body fluids, *e.g.*, blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the TLCC-2 nucleotide sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of SEQ ID NO:1 having a length of at least 20 bases, preferably at least 30 bases.

The TLCC-2 nucleotide sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such TLCC-2 probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, TLCC-2 primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

C. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining TLCC-2 protein and/or nucleic acid expression as well as TLCC-2 activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant or unwanted TLCC-2 expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with TLCC-2 protein, nucleic acid expression or activity. For example, mutations in a TLCC-2 gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby phophylactically treat an individual prior to the onset of a disorder characterized by or associated with TLCC-2 protein, nucleic acid expression or activity.

Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of TLCC-2 in clinical trials.

These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

An exemplary method for detecting the presence or absence of TLCC-2 protein or nucleic acid in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting TLCC-2 protein or nucleic acid (*e.g.*, mRNA, or genomic DNA) that encodes TLCC-2 protein such that the presence of TLCC-2 protein or nucleic acid is detected in the biological sample. A preferred agent for detecting TLCC-2 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to TLCC-2 mRNA or genomic DNA. The nucleic acid probe can be, for example, the TLCC-2 nucleic acid set forth in SEQ ID NO:1 or 3, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize

under stringent conditions to TLCC-2 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting TLCC-2 protein is an antibody capable of binding to TLCC-2 protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')2) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect TLCC-2 mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of TLCC-2 mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of TLCC-2 protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of TLCC-2 genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of TLCC-2 protein include introducing into a subject a labeled anti-TLCC-2 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting TLCC-2 protein, mRNA, or genomic DNA, such that the presence of TLCC-2 protein, mRNA or genomic DNA is detected in the

biological sample, and comparing the presence of TLCC-2 protein, mRNA or genomic DNA in the control sample with the presence of TLCC-2 protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of TLCC-2 in a biological sample. For example, the kit can comprise a labeled compound or agent capable of detecting TLCC-2 protein or mRNA in a biological sample; means for determining the amount of TLCC-2 in the sample; and means for comparing the amount of TLCC-2 in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect TLCC-2 protein or nucleic acid.

2. Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant or unwanted TLCC-2 expression or activity. As used herein, the term "aberrant" includes a TLCC-2 expression or activity which deviates from the wild type TLCC-2 expression or activity. Aberrant expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant TLCC-2 expression or activity is intended to include the cases in which a mutation in the TLCC-2 gene causes the TLCC-2 gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional TLCC-2 protein or a protein which does not function in a wild-type fashion, *e.g.*, a protein which does not interact with a TLCC-2 substrate, *e.g.*, a non-calcium channel subunit or ligand and/or a non-vanilloid receptor subunit or ligand, or one which interacts with a non-TLCC-2 substrate, *e.g.* a non-calcium channel subunit or ligand and/or a non-vanilloid receptor subunit or ligand. As used herein, the term "unwanted" includes an unwanted phenomenon involved in a biological response, such as cellular proliferation. For example, the term unwanted includes a TLCC-2 expression or activity which is undesirable in a subject.

The assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with a misregulation in TLCC-2 protein activity or nucleic acid expression, such as a CNS disorder (*e.g.*, a neurodegenerative disorder, a pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder). Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disorder associated with a misregulation in TLCC-2 protein activity or nucleic acid expression, such as a CNS disorder, a pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant or unwanted TLCC-2 expression or activity in which a test sample is obtained from a subject and TLCC-2 protein or nucleic acid (*e.g.*, mRNA or genomic DNA) is detected, wherein the presence of TLCC-2 protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted TLCC-2 expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, *e.g.,* neuronal cells, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted TLCC-2 expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a CNS disorder, pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant or unwanted TLCC-2 expression or activity in which a test sample is obtained and TLCC-2 protein or nucleic acid expression or activity is detected (*e.g.*, wherein the abundance of TLCC-2 protein or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant or unwanted TLCC-2 expression or activity).

The methods of the invention can also be used to detect genetic alterations in a TLCC-2 gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in TLCC-2 protein activity or nucleic acid expression, such as a CNS disorder, pain disorder, or a disorder of cellular growth, differentiation, or migration. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a TLCC-2 -protein, or the mis-expression of the TLCC-2 gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a TLCC-2 gene; 2) an addition of one or more nucleotides to a TLCC-2 gene; 3) a substitution of one or more nucleotides of a TLCC-2 gene, 4) a chromosomal rearrangement of a TLCC-2 gene; 5) an alteration in the level of a messenger RNA transcript of a TLCC-2 gene, 6) aberrant modification of a TLCC-2 gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a TLCC-2 gene, 8) a non-wild type level of a TLCC-2-protein, 9) allelic loss of a TLCC-2 gene, and 10) inappropriate post-translational modification of a TLCC-2-protein. As described herein, there are a large number of assays known in the art which can be used for detecting alterations in a TLCC-2 gene. A preferred biological sample is a tissue or serum sample isolated by conventional means from a subject.

In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the TLCC-2-gene (see Abravaya *et al.* (1995) *Nucleic Acids Res* .23:675-682). This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a TLCC-2 gene under conditions such that hybridization and amplification of the TLCC-2-gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of

the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.* (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a TLCC-2 gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in TLCC-2 can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in TLCC-2 can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, M.T. *et al. supra.* Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is

composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the TLCC-2 gene and detect mutations by comparing the sequence of the sample TLCC-2 with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the TLCC-2 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type TLCC-2 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in TLCC-2 cDNAs obtained from samples of cells. For example, the mutY enzyme of E. coli cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a TLCC-2 sequence, *e.g.*, a wild-type TLCC-2 sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in TLCC-2 genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA*: 86:2766, see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control TLCC-2 nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is

used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl Acad. Sci USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a TLCC-2 gene.

Furthermore, any cell type or tissue in which TLCC-2 is expressed may be utilized in the prognostic assays described herein.

3. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs) on the expression or activity of a TLCC-2 protein (*e.g.*, the modulation of membrane excitability) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase TLCC-2 gene expression, protein levels, or upregulate TLCC-2 activity, can be monitored in clinical trials of subjects exhibiting decreased TLCC-2 gene expression, protein levels, or downregulated TLCC-2 activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease TLCC-2 gene expression, protein levels, or downregulate TLCC-2 activity, can be monitored in clinical trials of subjects exhibiting increased TLCC-2 gene expression, protein levels, or upregulated TLCC-2 activity. In such clinical trials, the expression or activity of a TLCC-2 gene, and preferably, other genes that have been implicated in, for example, a TLCC-2-associated disorder can be used as a "read out" or markers of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including TLCC-2, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) which modulates TLCC-2 activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on TLCC-2-associated disorders (*e.g.*, disorders characterized by deregulated signaling or membrane excitation), for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of TLCC-2 and other genes implicated in the TLCC-2-associated disorder, respectively. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of TLCC-2 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, antibody, protein, peptide, nucleic acid, ribozyme, small

molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a TLCC-2 protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the TLCC-2 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the TLCC-2 protein, mRNA, or genomic DNA in the pre-administration sample with the TLCC-2 protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of TLCC-2 to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of TLCC-2 to lower levels than detected, *i.e.* to decrease the effectiveness of the agent. According to such an embodiment, TLCC-2 expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

D. Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant or unwanted TLCC-2 expression or activity, *e.g.* a CNS disorder, pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder.

"Treatment", or "treating" as used herein, is defined as the application or administration of a therapeutic agent to a patient, or application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disease or disorder, a symptom of disease or disorder or a predisposition toward a disease or disorder, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect the disease or disorder, the symptoms of the disease or disorder, or the predisposition toward disease. A therapeutic agent includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides.

With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how a patient's genes determine his or her response to a drug (*e.g.*, a patient's "drug response phenotype", or "drug response genotype"). Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the TLCC-2 molecules of the present invention or TLCC-2 modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant or unwanted TLCC-2 expression or activity, by administering to the subject a TLCC-2 or an agent which modulates TLCC-2 expression or at least one TLCC-2 activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted TLCC-2 expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the TLCC-2 aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of TLCC-2 aberrancy, for example, a TLCC-2, TLCC-2 agonist or TLCC-2 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating TLCC-2 expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a

TLCC-2 or agent that modulates one or more of the activities of TLCC-2 protein activity associated with the cell. An agent that modulates TLCC-2 protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a TLCC-2 protein (*e.g.*, a TLCC-2 substrate), a TLCC-2 antibody, a TLCC-2 agonist or antagonist, a peptidomimetic of a TLCC-2 agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more TLCC-2 activities. Examples of such stimulatory agents include active TLCC-2 protein and a nucleic acid molecule encoding TLCC-2 that has been introduced into the cell. In another embodiment, the agent inhibits one or more TLCC-2 activities. Examples of such inhibitory agents include antisense TLCC-2 nucleic acid molecules, ribozymes, anti-TLCC-2 antibodies, and TLCC-2 inhibitors. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a TLCC-2 protein or nucleic acid molecule, *e.g.*, a pain disorder. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) TLCC-2 expression or activity. In another embodiment, the method involves administering a TLCC-2 protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted TLCC-2 expression or activity.

Stimulation of TLCC-2 activity is desirable in situations in which TLCC-2 is abnormally downregulated and/or in which increased TLCC-2 activity is likely to have a beneficial effect. Likewise, inhibition of TLCC-2 activity is desirable in situations in which TLCC-2 is abnormally upregulated and/or in which decreased TLCC-2 activity is likely to have a beneficial effect, *e.g.*, in pain disorders.

3. Pharmacogenomics

The TLCC-2 molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on TLCC-2 activity (*e.g.*, TLCC-2 gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) TLCC-2-associated disorders

(*e.g.*, proliferative disorders) associated with aberrant or unwanted TLCC-2 activity. In conjunction with such treatment, pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a TLCC-2 molecule or TLCC-2 modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a TLCC-2 molecule or TLCC-2 modulator.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11): 983-985 and Linder, M.W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (*e.g.*, a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten-million known

single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drugs target is known (*e.g.*, a TLCC-2 protein of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (*e.g.*, a TLCC-2 molecule or TLCC-2 modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a TLCC-2 molecule or TLCC-2 modulator, such as a modulator identified by one of the exemplary screening assays described herein.

VI. Electronic Apparatus Readable Media and Arrays

Electronic apparatus readable media comprising TLCC-2 sequence information is also provided. As used herein, "TLCC-2 sequence information" refers to any nucleotide and/or amino acid sequence information particular to the TLCC-2 molecules of the present invention, including but not limited to full-length nucleotide and/or amino acid sequences, partial nucleotide and/or amino acid sequences, polymorphic sequences including single nucleotide polymorphisms (SNPs), epitope sequences, and the like. Moreover, information "related to" said TLCC-2 sequence information includes detection of the presence or absence of a sequence (*e.g.*, detection of expression of a sequence, fragment, polymorphism, etc.), determination of the level of a sequence (*e.g.*, detection of a level of expression, for example, a quantative detection), detection of a reactivity to a sequence (*e.g.*, detection of protein expression and/or levels, for example, using a sequence-specific antibody), and the like. As used herein, "electronic apparatus readable media" refers to any suitable medium for storing, holding or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like; general hard disks and hybrids of these categories such as magnetic/optical storage

media. The medium is adapted or configured for having recorded thereon TLCC-2 sequence information of the present invention.

As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the TLCC-2 sequence information.

A variety of software programs and formats can be used to store the sequence information on the electronic apparatus readable medium. For example, the sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of data processor structuring formats (*e.g.*, text file or database) may be employed in order to obtain or create a medium having recorded thereon the TLCC-2 sequence information.

By providing TLCC-2 sequence information in readable form, one can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the sequence information in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has a TLCC-2- associated disease or disorder or a pre-disposition to a TLCC-2-associated disease or disorder, wherein the method comprises the steps of determining TLCC-2 sequence information associated with the subject and based on the TLCC-2 sequence information, determining

whether the subject has a TLCC-2 -associated disease or disorder or a pre-disposition to a TLCC-2-associated disease or disorder and/or recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has a TLCC-2-associated disease or disorder or a pre-disposition to a disease associated with a TLCC-2 wherein the method comprises the steps of determining TLCC-2 sequence information associated with the subject, and based on the TLCC-2 sequence information, determining whether the subject has a TLCC-2 -associated disease or disorder or a pre-disposition to a TLCC-2 -associated disease or disorder, and/or recommending a particular treatment for the disease, disorder or pre-disease condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

The present invention also provides in a network, a method for determining whether a subject has a TLCC-2 -associated disease or disorder or a pre-disposition to a TLCC-2 -associated disease or disorder associated with TLCC-2, said method comprising the steps of receiving TLCC-2 sequence information from the subject and/or information related thereto, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to TLCC-2 and/or a [TLCC-2}-associated disease or disorder, and based on one or more of the phenotypic information, the TLCC-2 information (*e.g.*, sequence information and/or information related thereto), and the acquired information, determining whether the subject has a TLCC-2-associated disease or disorder or a pre-disposition to a TLCC-2-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention also provides a business method for determining whether a subject has a TLCC-2-associated disease or disorder or a pre-disposition to a TLCC-2-associated disease or disorder, said method comprising the steps of receiving information related to TLCC-2 (*e.g.*, sequence information and/or information related thereto), receiving phenotypic information associated with the subject, acquiring information from the network related to TLCC-2 and/or related to a TLCC-2-associated disease or disorder, and based on one or more of the phenotypic information, the TLCC-

2 information, and the acquired information, determining whether the subject has a TLCC-2-associated disease or disorder or a pre-disposition to a TLCC-2-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The invention also includes an array comprising a TLCC-2 sequence of the present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression, one of which can be TLCC-2. This allows a profile to be developed showing a battery of genes specifically expressed in one or more tissues.

In addition to such qualitative determination, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression *per se* and level of expression in that tissue. This is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of a TLCC-2-associated disease or disorder, progression of TLCC-2-associated disease or disorder, and processes, such a cellular transformation associated with the TLCC-2-associated disease or disorder.

The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells (*e.g.*, ascertaining the effect of TLCC-2 expression on the expression of other genes). This provides, for example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes (*e.g.*, including TLCC-2) that could serve as a molecular target for diagnosis or therapeutic intervention.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application, as well as the Figures and the Sequence Listing, are incorporated herein by reference.

EXAMPLES

EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF HUMAN TLCC-2 cDNA

In this example, the identification and characterization of the gene encoding human TLCC-2 (clone Fbh54420FL) is described.

Isolation of the TLCC-2 cDNA

The invention is based, at least in part, on the discovery of a human gene encoding a novel protein, referred to herein as TLCC-2. The entire sequence of the human clone Fbh54420FL was determined and found to contain an open reading frame termed human "TLCC-2". The nucleotide sequence of the human TLCC-2 gene is set forth in Figures 1 and in the Sequence Listing as SEQ ID NO:1 and 3. The amino acid sequence of the human TLCC-2 expression product is set forth in Figures 1 and in the Sequence Listing as SEQ ID NO: 2.

The nucleotide sequence encoding the human TLCC-2 protein is shown in Figure 1 and is set forth as SEQ ID NO:1. The protein encoded by this nucleic acid comprises about 580 amino acids and has the amino acid sequence shown in Figure 1 and set forth as SEQ ID NO:2. The coding region (open reading frame) of SEQ ID NO:1 is set forth as SEQ ID NO:3. Clone Fbh54420FL, comprising the coding region of human TLCC-2, was deposited with the American Type Culture Collection (ATCC®), 10801 University Boulevard, Manassas, VA 20110-2209, on _____, and assigned Accession No. _____.

Analysis of the Human TLCC-2 Molecules

A BLASTN 2.0 search against the dbEST database, using a score of 100 and a wordlength of 12 (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human TLCC-2 revealed that human IC54420 is 99% identical to tb24a12.x1 NCI_CGAP_kid12 *Homo sapiens* cDNA clone IMAGE:2055262 3' similar to WP:R13A5.1 CE01370 (Accession Number AI307240) over nucleotides 2077 to 1377. The search further revealed that human TLCC-2 is 98% identical to au44h03.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone IMAGE:2517653 3' similar to WP:R13A5.1 CE01370 (Accession Number AI816064) over nucleotides 2079-1375. This search further revealed that human TLCC-2 is 97% identical to wv36f01.x1 NCI_CGAP_Ov18 *Homo sapiens* cDNA clone IMAGE:2531641 3' similar to WP:R13A5.1 CE01370 (Accession Number AI951554) over nucleotides 2088 to 1407. This search further revealed that human TLCC-2 is 97% identical to wp80f10.x1 NCI_CGAP_Brn25 *Homo sapiens* cDNA clone IMAGE:2468107 3' similar to WP:R13A5.1 CE01370 (Accession Number AI942492) over nucleotides 2072-1422. The search further revealed that human TLCC-2 is 96% identical to nr72c11.s1 NCI_CGAP_Pr24 Homo sapiens cDNA clone IMAGE:1173524 similar to WP:R13A5.1 CE01370 (Accession Number AA641031) over nucleotides 2073-1407.

A BLASTX 2.0 search against the NRP/protot database, using a score of 100 and a wordlength of 3 (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403), of the translated nucleotide sequence of human TLCC-2 revealed that human TLCC-2 is 67% or less identical to fragments (*e.g.*, fragments of 185 amino acids or less) of unnamed protein product from *Homo sapiens* (Accession Number AK001868). The search further

revealed that human TLCC-2 is 27% or less identical to fragments (*e.g.*, fragments of 75 amino acids or less) of human polycystic kidney disease and receptor for egg jelly related protein (Accession Number AF116458). This search further revealed that human TLCC-2 is 43% or less identical to fragments (*e.g.*, fragments of 59 amino acids or less) of polycystic kidney disease and receptor for egg jelly related protein from *Mus musculus* (Accession Number AF116459). This search further revealed that human TLCC-2 is 70% or less identical to fragments (*e.g.*, fragments of 29 amino acids or less) of a 110 amino acid long hypothetical protein from *Aeropyrum pernix* (Accession Number AP000060) over the full length of this protein.

An alignment of the human TLCC-2 amino acid sequence with the amino acid sequences of vanilloid receptor 1 and vanilloid receptor 2 from *Homo sapiens* (Accession Numbers AAD26363 and AAD41724, SEQ ID NO:4 and SEQ ID NO:5, respectively) using the CLUSTAL W (1.74) multiple sequence alignment program is set forth in Figure 6. An alignment of the human TLCC-2 amino acid sequence with the amino acid sequence of polycystic kidney disease protein 2 from *Mus musculus* (Accession Number NP 032887, SEQ ID NO:6) using the CLUSTAL W(1.74) multiple sequence alignment program is set forth in Figure 7A-B. An alignment of the human TLCC-2 amino acid sequence with the amino acid sequence of human melastatin (Accession Number AAC8000, SEQ ID NO:7) using the CLUSTAL W(1.74) multiple sequence alignment program is set forth in Figure 8A-B.

A search was performed against the Memsat database (Figures 2 and 3), resulting in the identification of six transmembrane domains in the amino acid sequence of human TLCC-2 (SEQ ID NO:2) at about residues 70-86, 299-317, 354-371, 385-416, 428-447, and 497-521.

A search was also performed against the Prosite database (Figure 9) resulting in the identification of four N-glycosylation sites in the amino acid sequence of human TLCC-2 (SEQ ID NO:2) at about residues 159-162, 179-182, 220-223, and 230-233.

A search was also performed against the HMM database (Figure 4) resulting in the identification of a fibronectin type III domain in the amino acid sequence of human TLCC-2 (SEQ ID NO:2) at about residues 202-269 (score = 5).

A search was also performed against the ProDom database resulting in the identification of a 31K RNA-4 protein domain in the amino acid sequence of human TLCC-2 (SEQ ID NO:2) at about residues 397-443 (score = 73). The results of the search are set forth in Figure 5.

Tissue Distribution of Human TLCC-2 mRNA by Northern Analysis

This example describes the tissue distribution of TLCC-2 mRNA, as determined by Northern analysis.

Northern blot hybridizations with the various RNA samples are performed under standard conditions and washed under stringent conditions, *i.e.*, 0.2xSSC at 65°C. The DNA probe is radioactively labeled with ^{32}P-dCTP using the Prime-It kit (Stratagene, La Jolla, CA) according to the instructions of the supplier. Filters containing human mRNA (MultiTissue Northern I and MultiTissue Northern II from Clontech, Palo Alto, CA) are probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

Tissue Distribution of TLCC-2 mRNA by *In situ* Analysis

For *in situ* analysis, various tissues, *e.g.* tissues obtained from brain and spinal cord from monkey and rat, were first frozen on dry ice. Ten-micrometer-thick sections of the tissues were post-fixed with 4% formaldehyde in DEPC treated 1X phosphate-buffered saline at room temperature for 10 minutes before being rinsed twice in DEPC 1X phosphate-buffered saline and once in 0.1 M triethanolamine-HCl (pH 8.0). Following incubation in 0.25% acetic anhydride-0.1 M triethanolamine-HCl for 10 minutes, sections were rinsed in DEPC 2X SSC (1X SSC is 0.15M NaCl plus 0.015M sodium citrate). Tissue was then dehydrated through a series of ethanol washes, incubated in 100% chloroform for 5 minutes, and then rinsed in 100% ethanol for 1 minute and 95% ethanol for 1 minute and allowed to air dry.

Hybridizations were performed with ^{35}S-radiolabeled (5 X 10^7 cpm/ml) cRNA probes. Probes were incubated in the presence of a solution containing 600 mM NaCl, 10 mM Tris (pH 7.5), 1 mM EDTA, 0.01% sheared salmon sperm DNA, 0.01% yeast tRNA, 0.05% yeast total RNA type X1, 1X Denhardt's solution, 50% formamide, 10%

dextran sulfate, 100 mM dithiothreitol, 0.1% sodium dodecyl sulfate (SDS), and 0.1% sodium thiosulfate for 18 hours at 55°C.

After hybridization, slides were washed with 2X SSC. Sections are then sequentially incubated at 37°C in TNE (a solution containing 10 mM Tris-HCl (pH 7.6), 500 mM NaCl, and 1 mM EDTA), for 10 minutes, in TNE with 10µg of RNase A per ml for 30 minutes, and finally in TNE for 10 minutes. Slides were then rinsed with 2X SSC at room temperature, washed with 2X SSC at 50°C for 1 hour, washed with 0.2X SSC at 55°C for 1 hour, and 0.2X SSC at 60°C for 1 hour. Sections were then dehydrated rapidly through serial ethanol-0.3 M sodium acetate concentrations before being air dried and exposed to Kodak Biomax MR scientific imaging film for 24 hours and subsequently dipped in NB-2 photoemulsion and exposed at 4°C for 7 days before being developed and counter stained.

In situ hybridization results showed expression in monkey and rat brain (including cortex, stratium, and hippocampus), spinal cord, and dorsal root ganglia (DRG) neurons.

Is situ hybridization in rat animal models showed up-regulation of the TLCC-2 gene 10 days after unilateral chronic constriction injury (CCI). There was up-regulation of TLCC-2 seven days after axotomy and after intraplantar injection of complete Freund's adjuvant (CFA). These levels decreased to normal levels at later time points. No contralateral effects were observed. These results indicate that the TLCC-2 molecules of the present invention are up-regulated in response to painful simuli, and are therefore involved in nociception. Modulation, *e.g.*, inhibition, of expression or activity of the TLCC-2 molecules of the invention may therefore modulate nociception and provide treatment for pain disorders.

Tissue Expression Analysis of TLCC-2 mRNA Using Taqman Analysis

This example describes the tissue distribution of human TLCC-2 mRNA in a variety of cells and tissues, as determined using the TaqMan™ procedure. The Taqman™ procedure is a quantitative, reverse transcription PCR-based approach for detecting mRNA. The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold™ DNA Polymerase to cleave a TaqMan™ probe during PCR. Briefly, cDNA was generated from the samples of interest, *e.g.*, human brain, spinal cord, heart, kidney,

liver, lung, dorsal root ganglia, and skin, and used as the starting material for PCR amplification. In addition to the 5' and 3' gene-specific primers, a gene-specific oligonucleotide probe (complementary to the region being amplified) was included in the reaction (*i.e.*, the Taqman™ probe). The TaqMan™ probe includes the oligonucleotide with a fluorescent reporter dye covalently linked to the 5' end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N',N'-tetramethylrhodamine) at the 3' end of the probe.

During the PCR reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites. The 5'-3' nucleolytic activity of the AmpliTaq™ Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target. The probe fragments are then displaced from the target, and polymerization of the strand continues. The 3' end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product. RNA was prepared using the trizol method and treated with DNase to remove contaminating genomic DNA. cDNA was synthesized using standard techniques. Mock cDNA synthesis in the absence of reverse transcriptase resulted in samples with no detectable PCR amplification of the control gene confirms efficient removal of genomic DNA contamination.

Two human normal tissue panels indicated broad distribution of human TLCC-2 expression, with highest expression in human brain, followed by testis, placenta, adrenal gland, spinal cord, skin, and dorsal root ganglia (DRG) (See Figures 10 and 11).

EXAMPLE 2: EXPRESSION OF RECOMBINANT IC54420 PROTEIN IN BACTERIAL CELLS

In this example, TLCC-2 is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, TLCC-2 is fused to GST and this fusion polypeptide is expressed in *E. coli*, *e.g.*, strain PEB199. Expression of the GST-TLCC-2 fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

EXAMPLE 3: EXPRESSION OF RECOMBINANT IC54420 PROTEIN IN COS CELLS

To express the TLCC-2 gene in COS cells, the pcDNA/Amp vector by Invitrogen Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire TLCC-2 protein and an HA tag (Wilson *et al.* (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

To construct the plasmid, the TLCC-2 DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the TLCC-2 coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the TLCC-2 coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly,

MA). Preferably the two restriction sites chosen are different so that the TLCC-2 gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5□, SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the TLCC-2-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the IC54420 polypeptide is detected by radiolabelling (^{35}S-methionine or ^{35}S-cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labelled for 8 hours with ^{35}S-methionine (or ^{35}S-cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the TLCC-2 coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the TLCC-2 polypeptide is detected by radiolabelling and immunoprecipitation using a TLCC-2-specific monoclonal antibody.

EXAMPLE 4: REGULATION OF CALCIUM INFLUX THROUGH TLCC-2

This experiment describes the regulation of calcium influx though TLCC-2 in HEK293 cells as determined by Fluorometric Imaging Plate Reader experiments (FLIPR) (Molecular Devices Corp., Sunnyvale, CA).

The FLIPR is a screening tool for cell-based fluorescent assays which allows the simultaneous stimulation and measurement of separate cell populations in a high throughput format. Therefore, using this system, it is possible to quantify transient signals, such as the release of intracellular calcium, from cell populations, in parallel and in real time. The FLIPR contains chambers in which to hold the test plate and plates containing antagonists or agonists to be added to the test plate. The FLIPR utilizes an argon laser that provides discrete spectral lines spaced from approximately 350 to 530 nm. For use with fluorescent Ca^{2+} dyes, the 88-nm line of the laser is employed. The laser simultaneously illuminates the wells in a test plate. The image of each well in the plate is captured by a cooled charge coupled device (CCD) camera, which updates images once per second, if required, for the measurement of rapid calcium responses. Because both excitation and emission are read via the bottom of the plate, black-walled, transparent bottomed 96-well plates are used. Data captured by the CCD camera is converted to digital data and then transferred to a computer.

Briefly, a calcium indicator (*e.g.*, fluo-3/AM or Calcium Green-1/AM) was transferred to the culture medium. Because the FLIPR collects fluorescence from the bottom of the well, suspension cells require centrifugation to the base of the well following dye loading. Viable HEK293 cells were resuspended in loading medium and incubated for one hour. The cells were then centrifuged and resuspended with wash buffer. The cell suspension containing the dye was then aliquotted into each well of the black-walled, transparent bottomed 96-well plate and the plate was centrifuged. The FLIPR assay was then carried out and the results analyzed. (If adherent cells are used, they may be plated at an appropriate density in the 96-well plates and cultured overnight. Dye may then be loaded and incubated).

Results show a constitutive calcium influx through TLCC-2 in HEK293 cells that were incubated with NMDG/0 Ca^{+2} and stimulated afterwards with 5mM Ca^{+2}.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

What is claimed:

1. An isolated nucleic acid molecule selected from the group consisting of:

(a) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:1;

(b) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:3;

(c) a nucleic acid molecule comprising the nucleotide sequence contained in the plasmid deposited with ATCC® as Accession Number _____;

(d) a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3;

(e) a nucleic acid molecule comprising a fragment of at least 706 nucleotides of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3;

(f) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2;

(g) a nucleic acid molecule which encodes a polypeptide comprising an amino acid sequence at least about 60% identical to the amino acid sequence of SEQ ID NO:2;

(h) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 18 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2;

(i) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2;

(j) a nucleic acid molecule comprising a nucleotide sequence which is complementary to the nucleotide sequence of the nucleic acid molecule of any one of subparts (a) to (i); and

(k) a nucleic acid molecule which hybridizes to a complement of the nucleic acid molecule of any one of subparts (a) to (d), (f), (g) and (i) under stringent conditions.

2. An isolated nucleic acid molecule comprising the nucleic acid molecule of claim 1, and a nucleotide sequence encoding a heterologous polypeptide.

3. A vector comprising the nucleic acid molecule of claim 1.

4. The vector of claim 3, which is an expression vector.

5. A host cell transfected with the expression vector of claim 4.

6. A method of producing a polypeptide comprising culturing the host cell of claim 5 in an appropriate culture medium to, thereby, produce the polypeptide.

7. An isolated polypeptide selected from the group consisting of:

(a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 18 contiguous amino acids of SEQ ID NO:2;

(b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of SEQ ID NO:1 or SEQ ID NO:3 under stringent conditions;

(c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3;

(d) a polypeptide comprising an amino acid sequence which is at least 60% identical to the amino acid sequence of SEQ ID NO:2; and

(e) a polypeptide comprising the amino acid sequence of SEQ ID NO:2.

8. The polypeptide of claim 7, further comprising heterologous amino acid sequences.

9. An antibody which selectively binds to a polypeptide of claim 7.

10. A method for detecting the presence of a polypeptide of claim 7 in a sample comprising:

(a) contacting the sample with a compound which selectively binds to the polypeptide; and

(b) determining whether the compound binds to the polypeptide in the sample to thereby detect the presence of a polypeptide of claim 7 in the sample.

11. The method of claim 10, wherein the compound which binds to the polypeptide is an antibody.

12. A kit comprising a compound which selectively binds to a polypeptide of claim 7 and instructions for use.

13. A method for detecting the presence of a nucleic acid molecule of claim 1 in a sample comprising:

(a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and

(b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample to thereby detect the presence of a nucleic acid molecule of claim 1 in the sample.

14. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of claim 1 and instructions for use.

15. A method for identifying a compound which binds to or modulates the activity of a polypeptide of claim 7 comprising:

(a) contacting the polypeptide, or a cell expressing the polypeptide with a test compound; and

(b) determining whether the compound binds to or modulates the activity of the polypeptide, to thereby identify a compound which binds to or modulates the activity of the polypeptide.

16. A method of identifying a nucleic acid molecule associated with a pain disorder comprising:

(a) contacting a sample comprising nucleic acid molecules with a hybridization probe comprising at least 25 contiguous nucleotides of SEQ ID NO:1; and

(b) detecting the presence of a nucleic acid molecule in said sample that hybridizes to said probe, thereby identifying a nucleic acid molecule associated with a pain disorder.

17. A method of identifying a nucleic acid associated with a pain disorder comprising:

(a) contacting a sample comprising nucleic acid molecules with a first and a second amplification primer, said first primer comprising at least 25 contiguous nucleotides of SEQ ID NO:1 and said second primer comprising at least 25 contiguous nucleotides from the complement of SEQ ID NO:1;

(b) incubating said sample under conditions that allow nucleic acid amplification; and

(c) detecting the presence of a nucleic acid molecule in said sample that is amplified, thereby identifying a nucleic acid molecule associated with a pain disorder.

18. A method of identifying a polypeptide associated with a pain disorder comprising:

(a) contacting a sample comprising polypeptides with a TLCC-2 binding substance; and

(b) detecting the presence of a polypeptide in said sample that binds to said TLCC-2 binding substance, thereby identifying a polypeptide associated with a pain disorder.

19. The method of claim 18, wherein said binding substance is an antibody.

20. The method of claim 18, wherein said binding substance is detectably labeled.

21. A method for identifying a compound capable of treating a pain disorder comprising assaying the ability of the compound or agent to modulate TLCC-2 expression or activity, thereby identifying a compound capable of treating a pain disorder.

22. The method of claim 21, wherein said compound inhibits TLCC-2 expression or activity.

23. A method for identifying a subject suffering from a pain disorder comprising obtaining a biological sample from the subject, and detecting in the sample aberrant or abnormal TLCC-2 expression or activity, thereby identifying a subject suffering from a pain disorder.

24. A method for treating a subject having a pain disorder characterized by aberrant TLCC-2 polypeptide activity or aberrant TLCC-2 nucleic acid expression comprising administering to the subject a TLCC-2 modulator, thereby treating said subject having a pain disorder.

25. The method of claim 24, wherein the TLCC-2 modulator is selected from the group consisting of a small molecule, an antibody specific for TLCC-2, a TLCC-2 polypeptide, a fragment of a TLCC-2 polypeptide, a TLCC-2 nucleic acid molecule, a fragment of a TLCC-2 nucleic acid molecule, an antisense TLCC-2 nucleic acid molecule, and a ribozyme.

26. The method of claim 24, wherein said TLCC-2 modulator is administered in a pharmaceutically acceptable formulation.

27. The method of claim 24, wherein said TLCC-2 modulator is administered using a gene therapy vector.

28. A method for identifying a compound capable of treating a pain disorder characterized by aberrant TLCC-2 nucleic acid expression or TLCC-2 polypeptide activity comprising assaying the ability of the compound to modulate TLCC-2 nucleic acid expression or TLCC-2 polypeptide activity, thereby identifying a compound capable of treating a pain disorder characterized by aberrant TLCC-2 nucleic acid expression or TLCC-2 polypeptide activity.

29. A method for identifying a compound capable of modulating nociception comprising:

(a) contacting a cell with a test compound; and

(b) assaying the ability of the test compound to modulate the expression of a TLCC-2 nucleic acid or the activity of a TLCC-2 polypeptide;

thereby identifying a compound capable of modulating nociception.

30. The method of claim 29, wherein the TLCC-2 modulator is a TLCC-2 inhibitor.

Input file Fbh54420FL.seq; Output File 54420.trans
Sequence length 2095

```
GTCGACCCACGCGTCCGCGGACGCGTGGGCGGAGGGTTTGAAGCCGCGCCGCGAGGGAGCGAGGTCGCAGTGACAGCGG

                                                                M   T   A   P        4
CGGGCGATCGGACCCAGGCTGCCCCGCCGTACCCGCCTGCGTCCCGCGCTCCCGCCCCAGC ATG ACA GCC CCG       12

 A   G   P   R   G   S   E   T   E   R   L   L   T   P   N   P   G   Y   G   T      24
GCG GGT CCG CGC GGC TCA GAG ACC GAG CGG CTT CTG ACC CCC AAC CCC GGG TAT GGG ACC      72

 Q   A   G   P   S   P   A   P   P   T   P   P   E   E   E   D   L   R   R   R      44
CAG GCG GGG CCT TCA CCG GCC CCT CCG ACA CCC CCA GAA GAG GAA GAC CTT CGC CGT CGT     132

 L   K   Y   F   F   M   S   P   C   D   K   F   R   A   K   G   R   K   P   C      64
CTC AAA TAC TTT TTC ATG AGT CCC TGC GAC AAG TTT CGA GCC AAG GGC CGC AAG CCC TGC     192

 K   L   M   L   Q   V   V   K   I   L   V   V   T   V   Q   L   I   L   F   G      84
AAG CTG ATG CTG CAA GTG GTC AAG ATC CTG GTG GTC ACG GTG CAG CTC ATC CTG TTT GGG     252

 L   S   N   Q   L   A   V   T   F   R   E   E   N   T   I   A   F   R   H   L     104
CTC AGT AAT CAG CTG GCT GTG ACA TTC CGG GAA GAG AAC ACC ATC GCC TTC CGA CAC CTC     312

 F   L   L   G   Y   S   D   G   A   D   D   T   F   A   A   Y   T   R   E   Q     124
TTC CTG CTG GGC TAC TCG GAC GGA GCG GAT GAC ACC TTC GCA GCC TAC ACG CGG GAG CAG     372

 L   Y   Q   A   I   F   H   A   V   D   Q   Y   L   A   L   P   D   V   S   L     144
CTG TAC CAG GCC ATC TTC CAT GCT GTG GAC CAG TAC CTG GCG TTG CCT GAC GTG TCA CTG     432

 G   R   Y   A   Y   V   R   G   G   G   D   P   W   T   N   G   S   G   L   A     164
GGC CGG TAT GCG TAT GTC CGT GGT GGG GGT GAC CCT TGG ACC AAT GGC TCA GGG CTT GCT     492

 L   C   Q   R   Y   Y   H   R   G   H   V   D   P   A   N   D   T   F   D   I     184
CTC TGC CAG CGG TAC TAC CAC CGA GGC CAC GTG GAC CCG GCC AAC GAC ACA TTT GAC ATT     552

 D   P   M   V   V   T   D   C   I   Q   V   D   P   P   E   R   P   P   P   P     204
GAT CCG ATG GTG GTT ACT GAC TGC ATC CAG GTG GAT CCC CCC GAG CGG CCC CCT CCG CCC     612

 P   S   D   D   L   T   L   L   E   S   S   S   S   Y   K   N   L   T   L   K     224
CCC AGC GAC GAT CTC ACC CTC TTG GAA AGC AGC TCC AGT TAC AAG AAC CTC ACG CTC AAA     672

 F   H   K   L   V   N   V   T   I   H   F   R   L   K   T   I   N   L   Q   S     244
TTC CAC AAG CTG GTC AAT GTC ACC ATC CAC TTC CGG CTG AAG ACC ATT AAC CTC CAG AGC     732

 L   I   N   N   E   I   P   D   C   Y   T   F   S   V   L   I   T   F   D   N     264
CTC ATC AAT AAT GAG ATC CCG GAC TGC TAT ACC TTC AGC GTC CTG ATC ACG TTT GAC AAC     792

 K   A   H   S   G   R   I   P   I   S   L   E   T   Q   A   H   I   Q   E   C     284
AAA GCA CAC AGT GGG CGG ATC CCC ATC AGC CTG GAG ACC CAG GCC CAC ATC CAG GAG TGT     852

 K   H   P   S   V   F   Q   H   G   D   N   S   F   R   L   L   F   D   V   V     304
AAG CAC CCC AGT GTC TTC CAG CAC GGA GAC AAC AGC TTC CGG CTC CTG TTT GAC GTG GTG     912

 V   I   L   T   C   S   L   S   F   L   L   C   A   R   S   L   L   R   G   F     324
GTC ATC CTC ACC TGC TCC CTG TCC TTC CTC CTC TGC GCC CGC TCA CTC CTT CGA GGC TTC     972

 L   L   Q   N   E   F   V   G   F   M   W   R   Q   R   G   R   V   I   S   L     344
CTG CTG CAG AAC GAG TTT GTG GGG TTC ATG TGG CGG CAG CGG GGA CGG GTC ATC AGC CTG    1032

 W   E   R   L   E   F   V   N   G   W   Y   I   L   L   V   T   S   D   V   L     364
TGG GAG CGG CTG GAA TTT GTC AAT GGC TGG TAC ATC CTG CTC GTC ACC AGC GAT GTG CTC    1092
```

Fig. 1A

```
 T   I   S   G   T   I   M   K   I   G   I   E   A   K   N   L   A   S   Y   D    384
ACC ATC TCG GGC ACC ATC ATG AAG ATC GGC ATC GAG GCC AAG AAC TTG GCG AGC TAC GAC  1152

 V   C   S   I   L   L   G   T   S   T   L   L   V   W   V   G   V   I   R   Y    404
GTC TGC AGC ATC CTC CTG GGC ACC TCG ACG CTG CTG GTG TGG GTG GGC GTG ATC CGC TAC  1212

 L   T   F   F   H   N   Y   N   I   L   I   A   T   L   R   V   A   L   P   S    424
CTG ACC TTC TTC CAC AAC TAC AAT ATC CTC ATC GCC ACA CTG CGG GTG GCC CTG CCC AGC  1272

 V   M   R   F   C   C   C   V   A   V   I   Y   L   G   Y   C   F   C   G   W    444
GTC ATG CGC TTC TGC TGC TGC GTG GCT GTC ATC TAC CTG GGC TAC TGC TTC TGT GGC TGG  1332

 I   V   L   G   P   Y   H   V   K   F   R   S   L   S   M   V   S   E   C   L    464
ATC GTG CTG GGG CCC TAT CAT GTG AAG TTC CGC TCA CTC TCC ATG GTG TCT GAG TGC CTG  1392

 F   S   L   I   N   G   D   D   M   F   V   T   F   A   A   M   Q   A   Q   Q    484
TTC TCG CTC ATC AAT GGG GAC GAC ATG TTT GTG ACG TTC GCC GCC ATG CAG GCG CAG CAG  1452

 G   R   S   S   L   V   W   L   F   S   Q   L   Y   L   Y   S   F   I   S   L    504
GGC CGC AGC AGC CTG GTG TGG CTC TTC TCC CAG CTC TAC CTT TAC TCC TTC ATC AGC CTC  1512

 F   I   Y   M   V   L   S   L   F   I   A   L   I   T   G   A   Y   D   T   I    524
TTC ATC TAC ATG GTG CTC AGC CTC TTC ATC GCG CTC ATC ACC GGC GCC TAC GAC ACC ATC  1572

 K   H   P   G   G   A   G   A   E   E   S   E   L   Q   A   Y   I   A   Q   C    544
AAG CAT CCC GGC GGC GCA GGC GCA GAG GAG AGC GAG CTG CAG GCC TAC ATC GCA CAG TGC  1632

 Q   D   S   P   T   S   G   K   F   R   R   G   S   G   S   A   C   S   L   L    564
CAG GAC AGC CCC ACC TCC GGC AAG TTC CGC CGC GGG AGC GGC TCG GCC TGC AGC CTT CTC  1692

 C   C   C   G   R   D   P   S   E   E   H   S   L   L   V   N   *                581
TGC TGC TGC GGA AGG GAC CCC TCG GAG GAG CAT TCG CTG CTG GTG AAT TGA              1743

TTCGACCTGACTGCCGTTGGACCGTAGGCCCTGGACTGCAGAGACCCCCGCCCCGACCCCGCTTATTTATTTGTAGGGT

TTGCTTTTAAGGATCGGCTCCCTGTCGCGCCCGAGGAGGGCCTGGACCTTTCGTGTCGGACCCTTGGGGGCGGGGAGAC

TGGGTGGGAAGGTGTGAATAAAAGGGAAATTAAAAAAAAAAAAAAAATTTTAAA
```

Fig. 1B

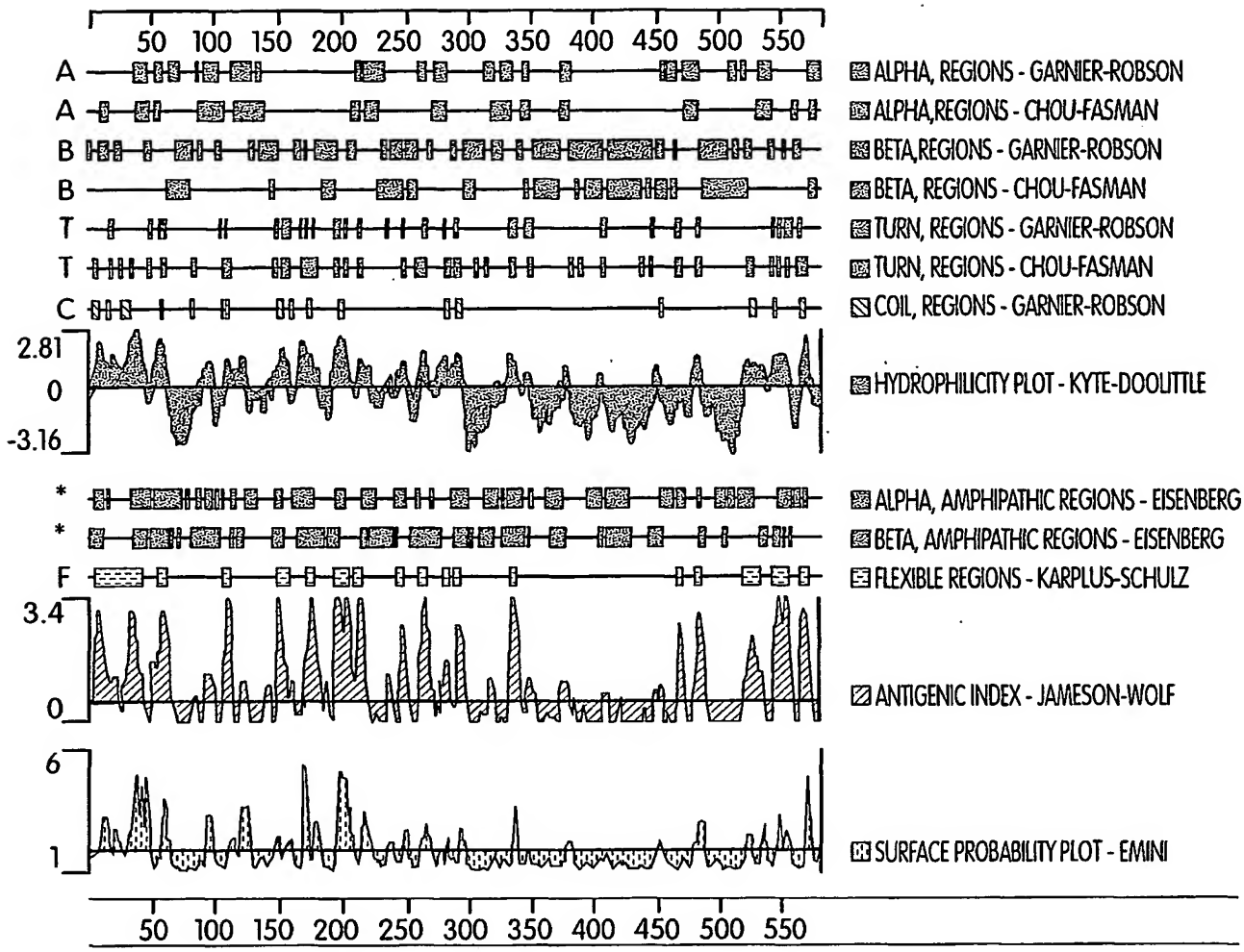


Fig. 2

TRANSMEMBRANE SEGMENTS PREDICTED BY MEMSAT

START	END	ORIENT	SCORE
70	86	ins-->out	2.0
299	317	out-->ins	3.7
354	371	ins-->out	2.4
385	402	out-->ins	3.1
428	447	ins-->out	4.7
497	521	out-->ins	6.4

>54420
MTAPAGPRGSETERLLTPNPGYGTQAGPSPAPPTPPEEEDLRRRLKYFFMSPCDKFRAKG
RKPCKLMLQVVKILVVTVQLILFGLSNQLAVTFREENTIAFRHLFLLGYSDGADDTFAAY
TREQLYQAIFHAVDQYLALPDVSLGRYAYVRGGGDPWTNGSGLALCQRYYHRGHVDPAND
TFDIDPMVVTDCIQVDPPERPPPPPSDDLTLLESSSSYKNLTLKFHKLVNVTIHFRLKTI
NLQSLINNEIPDCYTFSVLITFDNKAHSGRIPISLETQAHIQECKHPSVFQHGDNSFRLL
FDVVVILTCSLSFLLCARSLLRGFLLQNEFVGFMWRQRGRVISLWERLEFVNGWYILLVT
SDVLTISGTIMKIGIEAKNLASYDVCSILLGTSTLLVWVGVIRYLTFFHNYNILIATLRV
ALPSVMRFCCCVAVIYLGYCFCGWIVLGPYHVKFRSLSMVSECLFSLINGDDMFVTFAAM
QAQQGRSSLVWLFSQLYLYSFISLFIYMVLSLFIALITGAYDTIKHPGGAGAEESELQAY
IAQCQDSPTSGKFRRGSGSACSLLCCCGRDPSEEHSLLVN

Fig. 3

Protein Family / Domain Matches, HMMer Version 2

```
Searching for complete domains in PFAM
hmmpfam - search a single seq against HMM database
HMMER 2.1.1 (Dec 1998)
Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL).
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HMM file:                 /prod/ddm/seqanal/PFAM/pfam4.4/Pfam
Sequence file:            /tmp/orfanal.12622.aa
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
  Query:   54420

Scores for sequence family classification (score includes all domains):
Model    Description                                       Score    E-value  N
-------- -----------                                       -----    ------- ---
         [no hits above thresholds]

Parsed for domains:
Model    Domain  seq-f seq-t    hmm-f hmm-t      score  E-value
-------- ------- ----- -----    ----- -----      -----  -------
        [no hits above thresholds]

Alignments of top-scoring domains:
        [no hits above thresholds]
//
Searching for complete domains in SMART
hmmpfam - search a single seq against HMM database
HMMER 2.1.1 (Dec 1998)
Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL).
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HMM file:                 /ddm/robison/smart/smart/smart.all.hmms
Sequence file:            /tmp/orfanal.12622.aa
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Query:   54420

Scores for sequence family classification (score includes all domains):
Model    Description                                       Score    E-value  N
-------- -----------                                       -----    ------- ---
FN3_2                                                        5.0        9.4   1

Parsed for domains:
Model    Domain  seq-f seq-t    hmm-f hmm-t      score  E-value
-------- ------- ----- -----    ----- -----      -----  -------
FN3_2      1/1     202   269 ..     1    69 []     5.0      9.4

Alignments of top-scoring domains:
FN3_2: domain 1 of 1, from 202 to 269: score 5.0, E = 9.4
                    *->PgpPsnptelrvtdvtststsdltsvtlsWepP...ivgYr......
                       P+pPs++  1++ + +s++      +tl+  +  + ++++r ++ +
       54420   202     PPPPSDD--LTLLESSSSYK----NLTLKFHKLvnvTIHFRlktinl 242

                    .ytltgLkPgaepwteYefrVrAvngndaGeg<-*
                    +  + +  P+      Y+f+V    +n+a +g
       54420   243 qSLINNEIPD-----CYTFSVLITFDNKAHSG    269
```

Fig. 4

ProDom Matches

ProdomId	Start	End	Description	Score
View Prodom 35860	397	443	p99.2 (2) Q65670(1) Y32K(1) //PROTEIN 31K RNA-4	73
ProdomId	Start	End	Description	Score

View Prodom 35860

```
>35860 p99.2 (2) Q65670(1) Y32K(1)  // PROTEIN 31K RNA-4
       Length = 152

 Score = 73 (30.8 bits), Expect + 4.0, P = 0.98
 Identities = 15/50 (30%), Positives = 27/50 (54%)

Query:   397 VWVGVIRYLTFFHNYNILIATLRVALPSVMRFC-CCVAVIYLG--YCFCG 443
             VW+ ++   T +  +++++   +  LPS +  C CC  V Y    +CFCG
Sbjct:    21 VWILLLTSSTCYGYHDVVVDIEQCTLPSNIDGCVCCSGVCYFNDNHCFCG 70
```

Fig. 5

```
CLUSTAL W (1.74) multiple sequence alignment

humanVR1        MKKWSSTDLGTAADPLQKDTCPDPLDGDPNSRPPPAKPQLPTAKSRTRLFGKGDSEEAFP
hVR2            MTSPSSS-------PVFR---LETLDGGQEDGSEADRGKLD--------FGSG----LPP
54420.pro       MTAPAGP----------R--------GSETERLLTPNPGYG--------TQAG----PSP
                *.  :..          :        *.  .   . .                *      *

humanVR1        VDCPHEEGELDSCPTITVSPVITIQRPGDGPTGARLLSQDSVAASTEKTLRLYDRRSIFE
hVR2            MESQFQGEDRKFAPQIRVN---LNYRKGTG------------ASQPDPNR-FDRDRLFN
54420.pro       APPTPPEEE-DLRRRLKYF----------------------FMSPCDKFR--------A
                         : .    :                          *  .  *

humanVR1        AVAQNNCQDLESLLLFLQKSKKHLTDNEFKDPETGKTCLLKAMLNLHDGQNTTIPLLLEI
hVR2            AVSRGVPEDLAGLPEYLSKTSKYLTDSEYTEGSTGKTCLMKAVLNLKDGVNACILPLLQI
54420.pro       K-GRKPCKLMLQVVKILVVT---VQLILF-----GLSNQLAVTFREE---NTIAFRHLFL
                  .:   : :  :   *  :   :    :      * :  : . :. .   *:     * :

humanVR1        ARQTDSLKELVNASYTDSYYKGQTALHIAIERRNMALVTLLVENGADVQAAAHGDFFKKT
hVR2            DRDSGNPQPLVNAQCTDDYYRGHSALHIAIEKRSLQCVKLLVENGANVHARACGRFFQKG
54420.pro       LGYSDGADDTFAAYTREQLY--QAIFHAVDQYLALPDVSL----GRYAYVRGGGDPWTNG
                   :.. .  . *   :. *  :: :* . :   :  *.*    *  . . . *  : :

humanVR1        KGRPGFYFGELPLSLAACTNQLGIVKFLLQNSWQTADISARDSVGNTVLHALVEVADNTA
hVR2            QG-TCFYFGELPLSLAACTKQWDVVSYLLENPHQPASLQATDSQGNTVLHALVMISDNSA
54420.pro       SG--------L----ALC--QR----YYHRGHVDPAN-DTFDIDP-MVVTDCIQV-DPPE
                .*        *    * *  *     :  ..  :.*. .: *     *:   : : * .

humanVR1        DNTKFVTSMYNEILMLGAKLHPTLKLEELTNKKGMTPLALAAGTGKIGVLAYILQREIQE
hVR2            ENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKLAAKEGKIEIFRHILQREFSG
54420.pro       RPPPPPSDDLTLLESSSSYKNLTLKFHKLVNVTIHFRLKTIN----------LQSLINN
                      :.    :   .:    *:::..: *      *                **  :.

humanVR1        PECRHLSRKFTEWAYGPVHSSLYDLSCIDTCEKNSVLEVIAYSSSETPNRHDMLLVEPLN
hVR2            --LSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEIIAFH-CKSPHRHRMVVLEPLN
54420.pro       --E--IPD-----CY-----TFSVLITFDNKAHSGRIPISLET-----QAHIQECKHPS-
                    :.      .*     ::  *  .*.  ... : :          : *   . .*

humanVR1        RLLQDKWDRFVKRIFYFNFLVYCLYMIIFTMAAYYRP--VDGLPPFKMEKIGDYFRVTGE
hVR2            KLLQAKWDLLIPK-FFLNFLCNLIYMFIFTAVAYHQPTLKKQAAPHLKAEVGNSMLLTGH
54420.pro       -VFQH-GDNSFRLLFDVVVILTCSLSFLLCARSLLRG-----FL-LQNEFVGFMWRQRGR
                 ::*   *  .   * . .:      :::   :  :                :*     *.

humanVR1        ILSVLGGVYFFFRGIQYFLQRRPSMKTLFVDSYSEMLFFLQSLFMLATVVLYFSHLKEYV
hVR2            ILILLGGIYLLVGQLWYFWRRHVFIWISFIDSYFEILFLFQALLTVVSQVLCFLAIEWYL
54420.pro       VISLWERLEFVNG--WY---------ILLVTS--DVLTISGTIMKIGIEAKNLASYDVCS
                :: :   : :.     *           :: *  ::*    ::: :   .  :   .

humanVR1        ASMVFSLALGWTNMLYYTRGFQQMGIYAVMIEKMILRDLCRFMFVYIVFLFGFSTAVVTL
hVR2            PLLVSALVLGWLNLLYYTRGFQHTGIYSVMIQKVILRDLLRFLLIYLVFLFGFAVALVSL
54420.pro       ILLGTSTLLVWVGVIRYLTFFHNYNILIATLR-VALPSVMRFCCCVAVIYLG-------Y
                  :  :  * * .:: *   *:: .*  . :. : * .: **     *: :*

humanVR1        IEDGKNDSLPSESTSHRWRGPACRPPD----SSYNSLYSTCLELFKFTIGMGDLEFTENY
hVR2            SQEAWRPEAPTGPNATESVQPMEGQEDEGNGAQYRGILEASLELFKFTIGMGELAFQEQL
54420.pro       CFCGWIVLGPY----HVKFRSLSMVSE-----CLFSLINGDDMFVTFAA-MQAQQGRSSL
                   .     *          .    :        .: .    :..*:  *       ..

humanVR1        DFKAVFIILLLAYVILTYILLLNMLIALMGETVNKIAQESKNIWKLQRAITILDTEKSFL
hVR2            HFRGMVLLLLLAYVLLTYILLLNMLIALMSETVNSVATDSWSIWKLQKAISVLEMENGYW
54420.pro       -VWLFSQLYLYSFISLFIYMVLSLFIALITGAYDTIKHPG-GAGAEESELQAYIAQ----
                 .  .  : * ::: *   ::*.::***:  : :.:   . .    :  :     :

humanVR1        KCMRKAFRSGKLLQVGYTPDGKDDYRWCFRVDEVNWTTWNTNVGIINEDPGNCEGVKRTL
hVR2            WCRKKQ-RAGVMLTVGTKPDGSPDERWCFRVEEVNWASWEQTLPTLCEDP-SGAGVPRTL
54420.pro       -CQDSP--TSGKFRRG-S--GS--A--CSLLCCCGRDPSEEHSLLVN-------------
                 *  .   :.  :  * .  *.     *  :    .  . :     :

humanVR1        SFSLRSSRVSGRHWKNFALVPLLREASARDRQSAQPEEVYLRQFSGSLKPEDAEVFKSPA
hVR2            EN-------------------PVLASPPKEDEDGASEENYVPVQLLQSN------------
54420.pro       ------------------------------------------------------------

humanVR1        ASGEK
hVR2            -----
54420.pro       -----
```

Fig. 6

```
CLUSTAL W (1.74) multiple sequence alignment

54420.pro       ------------------------------------------------------------
mousePKD2       MVNSRRVQPQPPGDAGRSPAPRASGPGRLVAGGAGLAVPGGLGEQRGLEIEMERIRQAAA

54420.pro       ----------------------------------------------------------MTAPAG
mousePKD2       RDPPAGASASPSPPLSSCSRQAWSRDNPGFEAEEDDDDDEVEGEEGGMVVEMDVEWRPGS
                                                                         *..

54420.pro       PR---GSETERLLTPNPGYGTQAGPSPAPPTPPEEEDLRRRLKYFFMS--PCDK-FRAKG
mousePKD2       RRSASSSAVSSVGARGRGLGSYRGAAHLSGRRRRLEDQGAQCPSPAGGGDPLHRHLPLEG
                 *   .* .. : : . * *:  *.:  .    . **   :      . * .: :  :*

54420.pro       RKP-CKLMLQVVKIL--VVTVQLILFGLSNQ---LAVTFREENTIAFRHLFLLGYSDGAD
mousePKD2       QPPRVAWAERLVRGLRGLWGTRLMEESNANREKYLKSVLRELVTYLFFLVVLCILTYGMM
                : *      ::*: *  :  .:*:  . :*:   *  .:**  *  *  :.*   : *

54420.pro       DT-FAAYTR--EQLYQAIFHAVDQYLALPDVS--LGRYAYVRG---GGDPWTNGSG--LA
mousePKD2       SSNVYYYTRTLSQLFIDTPVSKTEKTNFKTLSSMEDFWKFTEGSFLDGLYWKAQTSNHTQ
                .: .  ***  .**:     :  :   :  :*   . : :..*   .*  *.  :.

54420.pro       LCQR---YYHRGHV-DP-------ANDTFDIDPMV---VTDCIQVD---PPERPPPPPSD
mousePKD2       ADNRSFIFYENLLLGVPRLRQLRVRNGSCSIPQDLRDEIKECYDVYSVSSEDRAPFGPRN
                  :*   :*..  :  *        *.: .*   :   :.:* :*    . :*.*  * :

54420.pro       DLTLLESSSSYKN--------LTLKFHKLVNVTIHFRLKTIN-LQSLINNEIPDCYTFSV
mousePKD2       GTAWMYTSEKELNGSSHWGIIASYSGAGYYLDLSRTREETAAQLAGLRRNFWLDRGTRAA
                . : : :*..  *          : .        : * :*   * .* .*   *  * :.

54420.pro       LITF--DNKAHSGRIPISLETQAHIQECKHPS-VFQHG-----DNSFRLLFDVVVILTCS
mousePKD2       FIDFSVYNANINLFCVVRLLAEFPATGGVVPSWQFQPVKLIRYVTAFDFFLAACEIIFCF
                :* *   *   .    : * ::  .       **  **        .:* ::: .  *: *

54420.pro       LSFLLCARSLLRGFLLQNEFVGFMWRQRGRVISLWERLEFVNGWYILLVTSDVLTISGTI
mousePKD2       FIIYYVVEEILEIRIHRLSYFRSFWNCLDVVIVVLSVVAMVINIYRMSNAEGLLQFLEDQ
                : :   ...:*.  : : .:.  :*.  . ** : . : :* . * :  :..:* :

54420.pro       MKIGIEAKNLASYDVCS-ILLGTSTLLVWVGVIRYLTFFHNYNILIATLRVALPSVMRFC
mousePKD2       NSFPN-FEHVAYWQIQFNNISAVMVFLVWIKLFKFINFNRTMSQLSTTMSRCAKDLFGFT
                 .:    :::* :::    : .. .:***: :::::.* :. . * :*:  .  .:: *

54420.pro       CCVAVIYLGYCFCGWIVLGPYHVKFRSLSMVSECLFSLINGDDMFVTFAAMQAQQGRSSL
mousePKD2       IMFSIIFLAYAQLAYLVFG---TQVDDFSTFQECIFTQFRIILGDINFAEIEEAN-R---
                  .::*:*.*.  .::*:*   .:. .:* ..**:*: :.     :.** ::  : *

54420.pro       VWLFSQLYLYSFISLFIYMVLSLFIALITGAYDTIKHPGGAGAEESELQAYIAQ-CQDSP
mousePKD2       --VLGPLYFTTFVFFMFFILLNMFLAIINDSYSEVKSDLAQQKAEMELSDLIRKGCQKAL
                  ::. **: :*: ::::::*.:*:*:*..:*. :*   .    * **.  * : **.:

54420.pro       TSGKFRRGSGSACSLLCCCG----------RDPSEE-HSLLVN----------------
mousePKD2       VKLKLKRNTVDAISESLRQGGGKLNFDELRQDLKGEGHTDAEIEAIFTKYDQDGDQELTE
                .. *::*.: .* *     *          :* . * *:

54420.pro       ------------------------------------------------------------
mousePKD2       REHQQMRDDLEKEREDLDLEHSSLPRPMSSRSFPRSLDDSEEEDDEDSGHSSRRRGSISS
```

Fig. 7A

```
54420.pro       ------------------------------------------------------------
mousePKD2       GVSYEEFQVLVRRVDRMEHSIGSIVSKIDAVIVKLEIMERAKLKRREVLGRLLDGVAEDA

54420.pro       ------------------------------------------------------------
mousePKD2       RLGRDSEIHREQMERLVREELERWESDDAASQTGHGVSTQVGLGGQPHPRNPRPPSSQSA

54420.pro       ----------------
mousePKD2       EGLEGGGGNGSANVHA
```

Fig. 7B

```
CLUSTAL W (1.74) multiple sequence alignment

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     MYIRVSYDTKPDSLLHLMVKDWQLELPKLLISVHGGLQNFEMQPKLKQVFGKGLIKAAMT

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     TGAWIFTGGVSTGVISHVGDALKDHSSKSRGRVCAIGIAPWGIVENKEDLVGKDVTRVYQ

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     TMSNPLSKLSVLNNSHTHFILADNGTLGKYGAEVKLRRLLEKHISLQKINTRLGQGVPLV

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     GLVVEGGPNVVSIVLEYLQEEPPIPVVICDGSGRASDILSFAHKYCEEGGIINESLREQL

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     LVTIQKTFNYNKAQSHQLFAIIMECMKKKELVTVFRMGSEGQQDIEMAILTALLKGTNVS

Fbh54420FL      -----MTAPAGPRGSET-ERLLTPNP---GYGTQAGPSPAPPT-----PPEE--EDLRRR
hMELASTATIN     APDQLSLALAWNRVDIARSQIFVFGPHWTPLGSLAPPTDSKATEKEKKPPMATTKGGRGK
                       * *  * . : .:::. .*     *: * *: : .*     **    :. * :

Fbh54420FL      LKYFFMSPCDKFRAKGRKPCKLMLQVVKILVVTVQLILFGLSNQLAVTFREENTIAFRHL
hMELASTATIN     GKGKKKGKVKEEVEEETDPRKIELLNWVNALEQAMLDALVLDRVDFVKLLIENGVNMQHF
                 *    .  .:   :  .* *: *     :  . *  : *..   *.:  ** : ::*:

Fbh54420FL      FLLG-----YSD--GADDTFAAYTR---EQLYQAIFHA--VDQYLALPDVSLGRY--AYV
hMELASTATIN     LTIPRLEELYNTRLGPPNTLHLLVRDVKKSNLPPDYHISLIDIGLVLEYLMGGAYRCNYT
                : :      *.   *. :*:   .*   :.   . :*   :*  *.*  :  * *   *.

Fbh54420FL      RGGG-DPWTNGSGLALCQRYYHRGHVD--PANDTFDIDPMVVTDCIQV---DPPERPPPP
hMELASTATIN     RKNFRTLYNNLFGPKRPKALKLLGMEDDEPPAKGKKKKKKKKEEEIDIDVDDPAVSRFQY
                * .    :.* *    :     *  *  *. .  . .     : *::   **.

Fbh54420FL      PSDDLTLLESSSSYKNLTLKFHKLVNVTIHFRLKTINLQSLINNEIPDCYTFS-VLITFD
hMELASTATIN     PFHELMVWAVLMKRQKMAVFLWQRGEESMAKALVACKLYKAMAHESSESDLVDDISQDLD
                * .:* :      . :::::: : :  : ::   * : :* . : :* .:.  .. :   :*

Fbh54420FL      N-KAHSGRIPISLETQ--AHIQECKHPSV---FQHGDNSFRLLFDVV------VILTCSL
hMELASTATIN     NNSKDFGQLALELLDQSYKHDEQIAMKLLTYELKNWSNSTCLKLAVAAKHRDFIAHTCSQ
                * . . *:.:.:.*  *   * ::     :   ::: .**  * : *.      :  ***

Fbh54420FL      SFL-------LCAR-----SLLRGFLLQNEFVGFMWRQRGRVISLWERLEFVNGWYILLV
hMELASTATIN     MLLTDMWMGRLRMRKNPGLKVIMGILLP-PTILFLEFRTYDDFSYQTSKENEDGKEKEEE
                 :*       *  *     .:: *:**    : *:  :    :*     *  :*

Fbh54420FL      TSDVLTISGTIMKIG-IEAKNLASY----DVCSILLGTSTLLVWVGVIRYLTFFHNYNIL
hMELASTATIN     NTDANADAGSRKGDEENEHKKQRSIPIGTKICEFYN-APIVKFWFYTISYLGYLLLFNYV
                .:*. : :*:       * *:  *     .:*.:   :. : .*. .* ** ::  :* :

Fbh54420FL      IATLRVALPSVMRFCCCVAVIYLGYCFCGWIVLGPYHVKFRSLSMVSECLFSLINGDDM-
hMELASTATIN     ILVRMDGWPSLQEWIVISYIVSLALEKIREILMSEPGKLSQKIKVWLQEYWNITDLVAIS
                * .   . **: .:     :: *.      *::.      :.:.:  :  :.: :   :
```

Fig. 8A

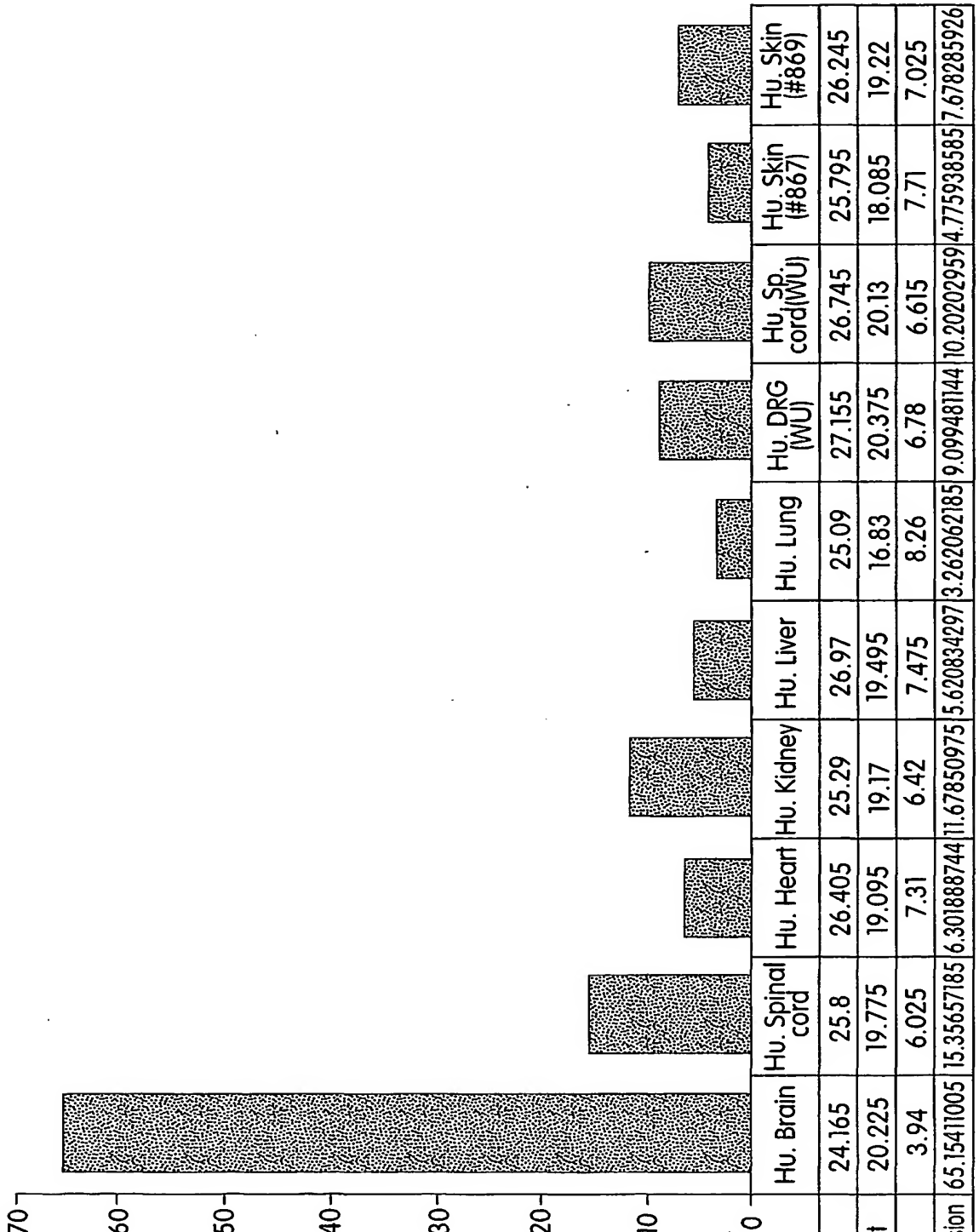
```
Fbh54420FL      -FVTFAAMQAQ-Q-----GRSS-----LVWLFSQLYLYSFISLF-------------IY
hMELASTATIN     TFMIGAILRLQNQPYMGYGRVIYCVDIIFWYIRVLDIFGVNKYLGPYVMMIGKMMIDMLY
                 *:  * :: * *     **       :.* :  * ::.. . :             :*

Fbh54420FL      MVLSLFIALITG--AYDTIKHP----GGAGAEESELQAY--IAQCQDSPTSGKFRRGSGS
hMELASTATIN     FVVIMLVVLMSFGVARQAILHPEEKPSWKLARNIFYMPYWMIYGEVFADQIDLYAMEINP
                :*: :::.*::   * ::* **    .   *.:    .*  *     :   . :    ..

Fbh54420FL      AC--SLLCCCG-RDPSEEHSLLVN---------------------------------
hMELASTATIN     PCGENLYDEEGKRLPPCIPGAWLTPALMACYLLVANILLVNLLIAVFNNTFFEVKSISNQ
                .*  .*    * * *.   .  :.

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     VWKFQRYQLIMTFHDRPVLPPPMIILSHIYIIIMRLSGRCRKKREGDQEERDRGLKLFLS

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     DEELKRLHEFEEQCVQEHFREKEDEQQSSSDERIRVTSERVENMSMRLEEINERETFMKT

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     SLQTVDLRLAQLEELSNRMVNALENLAGIDRSDLIQARSRASSECEATYLLRQSSINSAD

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     GYSLYRYHFNGEELLFEDTSLSTSPGTGVRKKTCSFRIKEEKDVKTHLVPECQNSLHLSL

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     GTSTSATPDGSHLAVDDLKNAEESKLGPDIGISKEDDERQTDSKKEETISPSLNKTDVIH

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     GQDKSDVQNTQLTVETTNIEGTISYPLEETKITRYFPDETINACKTMKSRSFVYSRGRKL

FbhS4420FL      ------------------------------------------------------------
hMELASTATIN     VGGVNQDVEYSSITDQQLTTEWQCQVQKITRSHSTDIPYIVSEAAVQAEQKEQFADMQDE

Fbh54420FL      ------------------------------------------------------------
hMELASTATIN     HHVAEAIPRIPRLSLTITDRNGMENLLSVKPDQTLGFPSLRSKSLHGHPRNVKSIQGKLD

Fbh54420FL      ---------------------------------
hMELASTATIN     RSGHASSVSSLVIVSGMTAEEKKVKKEKASTETEC
```

Fig. 8B

Prosite version: Release 12.2 of February 1995

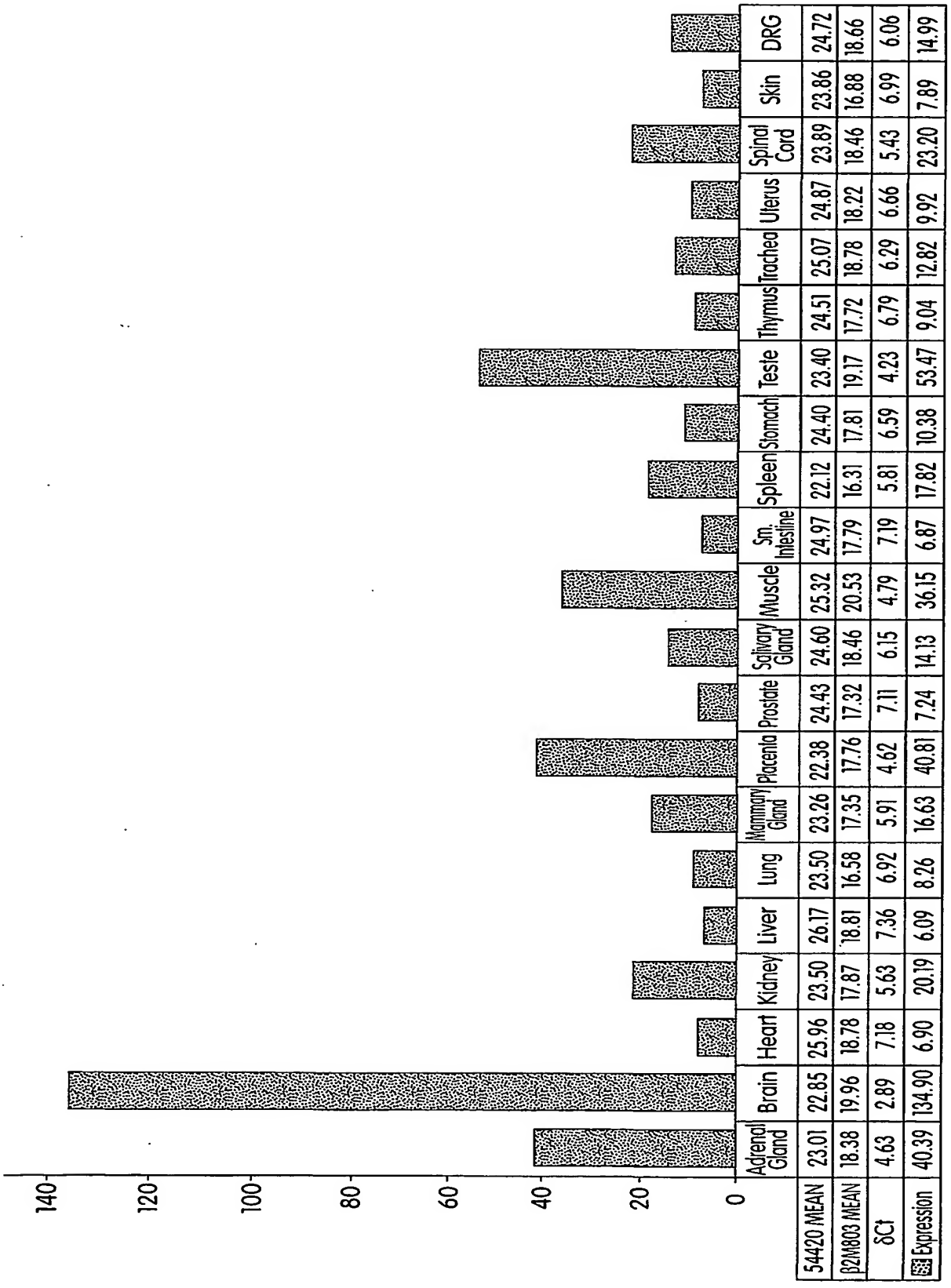
>PS00001|PDOC00001|ASN_GLYCOSYLATION N-glycosylation site.
Query: 159 NGSG 162
Query: 179 NDTF 182
Query: 220 NLTL 223
Ouery: 230 NVTI 233
>PS00004|PDOC00004|CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation site.
Ouery: 554 RRGS 557
>PS00005|PDOC00005|PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.
Query: 12 TER 14
Query: 92 TFR 94
Query: 217 SYK 219
Query: 222 TLK 224
Query: 268 SGR 270
Query: 296 SFR 298
Query: 417 TLR 419
Query: 523 TIK 525
Query: 550 SGK 552
>PS00006|PDOC00006|CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.
Query: 10 SETE 13
Query: 34 TPPE 37
Query: 51 SPCD 54
Query: 92 TFRE 95
Query: 210 TLLE 213
Query: 343 SLWE 346
>PS00008|PDOC00008|MYRISTYL N-myristoylation site.
Query: 112 GADDTF 117
Query: 160 GSGLAL 165
Query: 519 GAYDTI 524
Query: 528 GGAGAE 533
Query: 556 GSGSAC 561
>PS00009|PDOC00009|AMIDATION Amidation site.
Query: 59 KGRK 62
>PS00029|PDOC00029|LEUCINE_ZIPPER Leucine zipper pattern
Query: 300 LFDVVVILTCSLSFLLCARSLL 321
>PS00120|PDOC00110|LIPASE_SER Lipases, serine active site.
Query: 104 LFLLGYSDGA 113

Fig. 9



	Hu. Brain	Hu. Spinal cord	Hu. Heart	Hu. Kidney	Hu. Liver	Hu. Lung	Hu. DRG (WU)	Hu. Sp. cord(WU)	Hu. Skin (#867)	Hu. Skin (#869)
average Ct	24.165	25.8	26.405	25.29	26.97	25.09	27.155	26.745	25.795	26.245
average HK Ct	20.225	19.775	19.095	19.17	19.495	16.83	20.375	20.13	18.085	19.22
delta Ct	3.94	6.025	7.31	6.42	7.475	8.26	6.78	6.615	7.71	7.025
Relative Expression	65.15411005	15.35657185	6.301888744	11.67850975	5.620834297	3.262062185	9.099481144	10.20202959	4.775938585	7.678285926

Fig. 10



	Adrenal Gland	Brain	Heart	Kidney	Liver	Lung	Mammary Gland	Placenta	Prostate	Salivary Gland	Muscle	Sm. Intestine	Spleen	Stomach	Teste	Thymus	Trachea	Uterus	Spinal Cord	Skin	DRG
54420 MEAN	23.01	22.85	25.96	23.50	26.17	23.50	23.26	22.38	24.43	24.60	25.32	24.97	22.12	24.40	23.40	24.51	25.07	24.87	23.89	23.86	24.72
B2M803 MEAN	18.38	19.96	18.78	17.87	18.81	16.58	17.35	17.76	17.32	18.46	20.53	17.79	16.31	17.81	19.17	17.72	18.78	18.22	18.46	16.88	18.66
δCt	4.63	2.89	7.18	5.63	7.36	6.92	5.91	4.62	7.11	6.15	4.79	7.19	5.81	6.59	4.23	6.79	6.29	6.66	5.43	6.99	6.06
Expression	40.39	134.90	6.90	20.19	6.09	8.26	16.63	40.81	7.24	14.13	36.15	6.87	17.82	10.38	53.47	9.04	12.82	9.92	23.20	7.89	14.99

Fig. 11

SEQUENCE LISTING

<110> Millennium Pharmaceuticals, Inc.

<120> 54420, A NOVEL HUMAN CALCIUM CHANNEL

<130> MNI-125CPPC

<140>
<141>

<150> US 09/544,797
<151> 2000-04-07

<160> 7

<170> PatentIn Ver. 2.0

<210> 1
<211> 2095
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (141)..(1880)

<400> 1
gtcgacccac gcgtccgcgg acgcgtgggc ggagggtttg aagccgcgcc gcgagggagc 60

gaggtcgcag tgacagcggc gggcgatcgg acccaggctg ccccgccgta cccgcctgcg 120

tcccgcgctc ccgccccagc atg aca gcc ccg gcg ggt ccg cgc ggc tca gag 173
Met Thr Ala Pro Ala Gly Pro Arg Gly Ser Glu
1 5 10

acc gag cgg ctt ctg acc ccc aac ccc ggg tat ggg acc cag gcg ggg 221
Thr Glu Arg Leu Leu Thr Pro Asn Pro Gly Tyr Gly Thr Gln Ala Gly
15 20 25

cct tca ccg gcc cct ccg aca ccc cca gaa gag gaa gac ctt cgc cgt 269
Pro Ser Pro Ala Pro Pro Thr Pro Pro Glu Glu Glu Asp Leu Arg Arg
30 35 40

cgt ctc aaa tac ttt ttc atg agt ccc tgc gac aag ttt cga gcc aag 317
Arg Leu Lys Tyr Phe Phe Met Ser Pro Cys Asp Lys Phe Arg Ala Lys
45 50 55

ggc cgc aag ccc tgc aag ctg atg ctg caa gtg gtc aag atc ctg gtg 365
Gly Arg Lys Pro Cys Lys Leu Met Leu Gln Val Val Lys Ile Leu Val
60 65 70 75

gtc acg gtg cag ctc atc ctg ttt ggg ctc agt aat cag ctg gct gtg 413
Val Thr Val Gln Leu Ile Leu Phe Gly Leu Ser Asn Gln Leu Ala Val
80 85 90

aca ttc cgg gaa gag aac acc atc gcc ttc cga cac ctc ttc ctg ctg 461
Thr Phe Arg Glu Glu Asn Thr Ile Ala Phe Arg His Leu Phe Leu Leu
95 100 105

ggc tac tcg gac gga gcg gat gac acc ttc gca gcc tac acg cgg gag 509

```
Gly Tyr Ser Asp Gly Ala Asp Asp Thr Phe Ala Ala Tyr Thr Arg Glu
        110                 115                 120

cag ctg tac cag gcc atc ttc cat gct gtg gac cag tac ctg gcg ttg   557
Gln Leu Tyr Gln Ala Ile Phe His Ala Val Asp Gln Tyr Leu Ala Leu
    125                 130                 135

cct gac gtg tca ctg ggc cgg tat gcg tat gtc cgt ggt ggg ggt gac   605
Pro Asp Val Ser Leu Gly Arg Tyr Ala Tyr Val Arg Gly Gly Gly Asp
140                 145                 150                 155

cct tgg acc aat ggc tca ggg ctt gct ctc tgc cag cgg tac tac cac   653
Pro Trp Thr Asn Gly Ser Gly Leu Ala Leu Cys Gln Arg Tyr Tyr His
                160                 165                 170

cga ggc cac gtg gac ccg gcc aac gac aca ttt gac att gat ccg atg   701
Arg Gly His Val Asp Pro Ala Asn Asp Thr Phe Asp Ile Asp Pro Met
            175                 180                 185

gtg gtt act gac tgc atc cag gtg gat ccc ccc gag cgg ccc cct ccg   749
Val Val Thr Asp Cys Ile Gln Val Asp Pro Pro Glu Arg Pro Pro Pro
        190                 195                 200

ccc ccc agc gac gat ctc acc ctc ttg gaa agc agc tcc agt tac aag   797
Pro Pro Ser Asp Asp Leu Thr Leu Leu Glu Ser Ser Ser Ser Tyr Lys
    205                 210                 215

aac ctc acg ctc aaa ttc cac aag ctg gtc aat gtc acc atc cac ttc   845
Asn Leu Thr Leu Lys Phe His Lys Leu Val Asn Val Thr Ile His Phe
220                 225                 230                 235

cgg ctg aag acc att aac ctc cag agc ctc atc aat aat gag atc ccg   893
Arg Leu Lys Thr Ile Asn Leu Gln Ser Leu Ile Asn Asn Glu Ile Pro
                240                 245                 250

gac tgc tat acc ttc agc gtc ctg atc acg ttt gac aac aaa gca cac   941
Asp Cys Tyr Thr Phe Ser Val Leu Ile Thr Phe Asp Asn Lys Ala His
            255                 260                 265

agt ggg cgg atc ccc atc agc ctg gag acc cag gcc cac atc cag gag   989
Ser Gly Arg Ile Pro Ile Ser Leu Glu Thr Gln Ala His Ile Gln Glu
        270                 275                 280

tgt aag cac ccc agt gtc ttc cag cac gga gac aac agc ttc cgg ctc   1037
Cys Lys His Pro Ser Val Phe Gln His Gly Asp Asn Ser Phe Arg Leu
    285                 290                 295

ctg ttt gac gtg gtg gtc atc ctc acc tgc tcc ctg tcc ttc ctc ctc   1085
Leu Phe Asp Val Val Val Ile Leu Thr Cys Ser Leu Ser Phe Leu Leu
300                 305                 310                 315

tgc gcc cgc tca ctc ctt cga ggc ttc ctg ctg cag aac gag ttt gtg   1133
Cys Ala Arg Ser Leu Leu Arg Gly Phe Leu Leu Gln Asn Glu Phe Val
                320                 325                 330

ggg ttc atg tgg cgg cag cgg gga cgg gtc atc agc ctg tgg gag cgg   1181
Gly Phe Met Trp Arg Gln Arg Gly Arg Val Ile Ser Leu Trp Glu Arg
            335                 340                 345

ctg gaa ttt gtc aat ggc tgg tac atc ctg ctc gtc acc agc gat gtg   1229
Leu Glu Phe Val Asn Gly Trp Tyr Ile Leu Leu Val Thr Ser Asp Val
```

```
        350                 355                 360

ctc acc atc tcg ggc acc atc atg aag atc ggc atc gag gcc aag aac   1277
Leu Thr Ile Ser Gly Thr Ile Met Lys Ile Gly Ile Glu Ala Lys Asn
    365                 370                 375

ttg gcg agc tac gac gtc tgc agc atc ctc ctg ggc acc tcg acg ctg   1325
Leu Ala Ser Tyr Asp Val Cys Ser Ile Leu Leu Gly Thr Ser Thr Leu
380                 385                 390                 395

ctg gtg tgg gtg ggc gtg atc cgc tac ctg acc ttc ttc cac aac tac   1373
Leu Val Trp Val Gly Val Ile Arg Tyr Leu Thr Phe Phe His Asn Tyr
                400                 405                 410

aat atc ctc atc gcc aca ctg cgg gtg gcc ctg ccc agc gtc atg cgc   1421
Asn Ile Leu Ile Ala Thr Leu Arg Val Ala Leu Pro Ser Val Met Arg
            415                 420                 425

ttc tgc tgc tgc gtg gct gtc atc tac ctg ggc tac tgc ttc tgt ggc   1469
Phe Cys Cys Cys Val Ala Val Ile Tyr Leu Gly Tyr Cys Phe Cys Gly
        430                 435                 440

tgg atc gtg ctg ggg ccc tat cat gtg aag ttc cgc tca ctc tcc atg   1517
Trp Ile Val Leu Gly Pro Tyr His Val Lys Phe Arg Ser Leu Ser Met
    445                 450                 455

gtg tct gag tgc ctg ttc tcg ctc atc aat ggg gac gac atg ttt gtg   1565
Val Ser Glu Cys Leu Phe Ser Leu Ile Asn Gly Asp Asp Met Phe Val
460                 465                 470                 475

acg ttc gcc gcc atg cag gcg cag cag ggc cgc agc agc ctg gtg tgg   1613
Thr Phe Ala Ala Met Gln Ala Gln Gln Gly Arg Ser Ser Leu Val Trp
                480                 485                 490

ctc ttc tcc cag ctc tac ctt tac tcc ttc atc agc ctc ttc atc tac   1661
Leu Phe Ser Gln Leu Tyr Leu Tyr Ser Phe Ile Ser Leu Phe Ile Tyr
            495                 500                 505

atg gtg ctc agc ctc ttc atc gcg ctc atc acc ggc gcc tac gac acc   1709
Met Val Leu Ser Leu Phe Ile Ala Leu Ile Thr Gly Ala Tyr Asp Thr
        510                 515                 520

atc aag cat ccc ggc ggc gca ggc gca gag gag agc gag ctg cag gcc   1757
Ile Lys His Pro Gly Gly Ala Gly Ala Glu Glu Ser Glu Leu Gln Ala
    525                 530                 535

tac atc gca cag tgc cag gac agc ccc acc tcc ggc aag ttc cgc cgc   1805
Tyr Ile Ala Gln Cys Gln Asp Ser Pro Thr Ser Gly Lys Phe Arg Arg
540                 545                 550                 555

ggg agc ggc tcg gcc tgc agc ctt ctc tgc tgc tgc gga agg gac ccc   1853
Gly Ser Gly Ser Ala Cys Ser Leu Leu Cys Cys Cys Gly Arg Asp Pro
                560                 565                 570

tcg gag gag cat tcg ctg ctg gtg aat tgattcgacc tgactgccgt         1900
Ser Glu Glu His Ser Leu Leu Val Asn
            575                 580

tggaccgtag gccctggact gcagagaccc ccgccccgac cccgcttatt tatttgtagg 1960

gtttgctttt aaggatcggc tccctgtcgc gcccgaggag ggcctggacc tttcgtgtcg 2020
```

gacccttggg ggcggggaga ctgggtggga aggtgtgaat aaaagggaaa ttaaaaaaaa 2080

aaaaaaaatt ttaaa 2095

<210> 2
<211> 580
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Thr Ala Pro Ala Gly Pro Arg Gly Ser Glu Thr Glu Arg Leu Leu
  1               5                  10                  15

Thr Pro Asn Pro Gly Tyr Gly Thr Gln Ala Gly Pro Ser Pro Ala Pro
             20                  25                  30

Pro Thr Pro Pro Glu Glu Glu Asp Leu Arg Arg Arg Leu Lys Tyr Phe
         35                  40                  45

Phe Met Ser Pro Cys Asp Lys Phe Arg Ala Lys Gly Arg Lys Pro Cys
     50                  55                  60

Lys Leu Met Leu Gln Val Val Lys Ile Leu Val Val Thr Val Gln Leu
 65                  70                  75                  80

Ile Leu Phe Gly Leu Ser Asn Gln Leu Ala Val Thr Phe Arg Glu Glu
                 85                  90                  95

Asn Thr Ile Ala Phe Arg His Leu Phe Leu Leu Gly Tyr Ser Asp Gly
            100                 105                 110

Ala Asp Asp Thr Phe Ala Ala Tyr Thr Arg Glu Gln Leu Tyr Gln Ala
        115                 120                 125

Ile Phe His Ala Val Asp Gln Tyr Leu Ala Leu Pro Asp Val Ser Leu
    130                 135                 140

Gly Arg Tyr Ala Tyr Val Arg Gly Gly Gly Asp Pro Trp Thr Asn Gly
145                 150                 155                 160

Ser Gly Leu Ala Leu Cys Gln Arg Tyr Tyr His Arg Gly His Val Asp
                165                 170                 175

Pro Ala Asn Asp Thr Phe Asp Ile Asp Pro Met Val Val Thr Asp Cys
            180                 185                 190

Ile Gln Val Asp Pro Pro Glu Arg Pro Pro Pro Pro Pro Ser Asp Asp
        195                 200                 205

Leu Thr Leu Leu Glu Ser Ser Ser Ser Tyr Lys Asn Leu Thr Leu Lys
    210                 215                 220

Phe His Lys Leu Val Asn Val Thr Ile His Phe Arg Leu Lys Thr Ile
225                 230                 235                 240

Asn Leu Gln Ser Leu Ile Asn Asn Glu Ile Pro Asp Cys Tyr Thr Phe
                245                 250                 255

Ser Val Leu Ile Thr Phe Asp Asn Lys Ala His Ser Gly Arg Ile Pro
```

```
            260                 265                 270

Ile Ser Leu Glu Thr Gln Ala His Ile Gln Glu Cys Lys His Pro Ser
        275                 280                 285

Val Phe Gln His Gly Asp Asn Ser Phe Arg Leu Leu Phe Asp Val Val
    290                 295                 300

Val Ile Leu Thr Cys Ser Leu Ser Phe Leu Leu Cys Ala Arg Ser Leu
305                 310                 315                 320

Leu Arg Gly Phe Leu Leu Gln Asn Glu Phe Val Gly Phe Met Trp Arg
                325                 330                 335

Gln Arg Gly Arg Val Ile Ser Leu Trp Glu Arg Leu Glu Phe Val Asn
            340                 345                 350

Gly Trp Tyr Ile Leu Leu Val Thr Ser Asp Val Leu Thr Ile Ser Gly
        355                 360                 365

Thr Ile Met Lys Ile Gly Ile Glu Ala Lys Asn Leu Ala Ser Tyr Asp
    370                 375                 380

Val Cys Ser Ile Leu Leu Gly Thr Ser Thr Leu Leu Val Trp Val Gly
385                 390                 395                 400

Val Ile Arg Tyr Leu Thr Phe Phe His Asn Tyr Asn Ile Leu Ile Ala
                405                 410                 415

Thr Leu Arg Val Ala Leu Pro Ser Val Met Arg Phe Cys Cys Cys Val
            420                 425                 430

Ala Val Ile Tyr Leu Gly Tyr Cys Phe Cys Gly Trp Ile Val Leu Gly
        435                 440                 445

Pro Tyr His Val Lys Phe Arg Ser Leu Ser Met Val Ser Glu Cys Leu
    450                 455                 460

Phe Ser Leu Ile Asn Gly Asp Asp Met Phe Val Thr Phe Ala Ala Met
465                 470                 475                 480

Gln Ala Gln Gln Gly Arg Ser Ser Leu Val Trp Leu Phe Ser Gln Leu
                485                 490                 495

Tyr Leu Tyr Ser Phe Ile Ser Leu Phe Ile Tyr Met Val Leu Ser Leu
            500                 505                 510

Phe Ile Ala Leu Ile Thr Gly Ala Tyr Asp Thr Ile Lys His Pro Gly
        515                 520                 525

Gly Ala Gly Ala Glu Glu Ser Glu Leu Gln Ala Tyr Ile Ala Gln Cys
    530                 535                 540

Gln Asp Ser Pro Thr Ser Gly Lys Phe Arg Arg Gly Ser Gly Ser Ala
545                 550                 555                 560

Cys Ser Leu Leu Cys Cys Cys Gly Arg Asp Pro Ser Glu Glu His Ser
                565                 570                 575

Leu Leu Val Asn
            580
```

<210> 3
<211> 1740
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1740)

<400> 3

```
atg aca gcc ccg gcg ggt ccg cgc ggc tca gag acc gag cgg ctt ctg   48
Met Thr Ala Pro Ala Gly Pro Arg Gly Ser Glu Thr Glu Arg Leu Leu
  1               5                  10                  15

acc ccc aac ccc ggg tat ggg acc cag gcg ggg cct tca ccg gcc cct   96
Thr Pro Asn Pro Gly Tyr Gly Thr Gln Ala Gly Pro Ser Pro Ala Pro
            20                  25                  30

ccg aca ccc cca gaa gag gaa gac ctt cgc cgt cgt ctc aaa tac ttt   144
Pro Thr Pro Pro Glu Glu Glu Asp Leu Arg Arg Arg Leu Lys Tyr Phe
        35                  40                  45

ttc atg agt ccc tgc gac aag ttt cga gcc aag ggc cgc aag ccc tgc   192
Phe Met Ser Pro Cys Asp Lys Phe Arg Ala Lys Gly Arg Lys Pro Cys
    50                  55                  60

aag ctg atg ctg caa gtg gtc aag atc ctg gtg gtc acg gtg cag ctc   240
Lys Leu Met Leu Gln Val Val Lys Ile Leu Val Val Thr Val Gln Leu
65                  70                  75                  80

atc ctg ttt ggg ctc agt aat cag ctg gct gtg aca ttc cgg gaa gag   288
Ile Leu Phe Gly Leu Ser Asn Gln Leu Ala Val Thr Phe Arg Glu Glu
                85                  90                  95

aac acc atc gcc ttc cga cac ctc ttc ctg ctg ggc tac tcg gac gga   336
Asn Thr Ile Ala Phe Arg His Leu Phe Leu Leu Gly Tyr Ser Asp Gly
            100                 105                 110

gcg gat gac acc ttc gca gcc tac acg cgg gag cag ctg tac cag gcc   384
Ala Asp Asp Thr Phe Ala Ala Tyr Thr Arg Glu Gln Leu Tyr Gln Ala
        115                 120                 125

atc ttc cat gct gtg gac cag tac ctg gcg ttg cct gac gtg tca ctg   432
Ile Phe His Ala Val Asp Gln Tyr Leu Ala Leu Pro Asp Val Ser Leu
    130                 135                 140

ggc cgg tat gcg tat gtc cgt ggt ggg ggt gac cct tgg acc aat ggc   480
Gly Arg Tyr Ala Tyr Val Arg Gly Gly Gly Asp Pro Trp Thr Asn Gly
145                 150                 155                 160

tca ggg ctt gct ctc tgc cag cgg tac tac cac cga ggc cac gtg gac   528
Ser Gly Leu Ala Leu Cys Gln Arg Tyr Tyr His Arg Gly His Val Asp
                165                 170                 175

ccg gcc aac gac aca ttt gac att gat ccg atg gtg gtt act gac tgc   576
Pro Ala Asn Asp Thr Phe Asp Ile Asp Pro Met Val Val Thr Asp Cys
            180                 185                 190

atc cag gtg gat ccc ccc gag cgg ccc cct ccg ccc ccc agc gac gat   624
```

```
Ile Gln Val Asp Pro Pro Glu Arg Pro Pro Pro Pro Pro Ser Asp Asp
        195                 200                 205

ctc acc ctc ttg gaa agc agc tcc agt tac aag aac ctc acg ctc aaa   672
Leu Thr Leu Leu Glu Ser Ser Ser Ser Tyr Lys Asn Leu Thr Leu Lys
    210                 215                 220

ttc cac aag ctg gtc aat gtc acc atc cac ttc cgg ctg aag acc att   720
Phe His Lys Leu Val Asn Val Thr Ile His Phe Arg Leu Lys Thr Ile
225                 230                 235                 240

aac ctc cag agc ctc atc aat aat gag atc ccg gac tgc tat acc ttc   768
Asn Leu Gln Ser Leu Ile Asn Asn Glu Ile Pro Asp Cys Tyr Thr Phe
                245                 250                 255

agc gtc ctg atc acg ttt gac aac aaa gca cac agt ggg cgg atc ccc   816
Ser Val Leu Ile Thr Phe Asp Asn Lys Ala His Ser Gly Arg Ile Pro
            260                 265                 270

atc agc ctg gag acc cag gcc cac atc cag gag tgt aag cac ccc agt   864
Ile Ser Leu Glu Thr Gln Ala His Ile Gln Glu Cys Lys His Pro Ser
        275                 280                 285

gtc ttc cag cac gga gac aac agc ttc cgg ctc ctg ttt gac gtg gtg   912
Val Phe Gln His Gly Asp Asn Ser Phe Arg Leu Leu Phe Asp Val Val
    290                 295                 300

gtc atc ctc acc tgc tcc ctg tcc ttc ctc ctc tgc gcc cgc tca ctc   960
Val Ile Leu Thr Cys Ser Leu Ser Phe Leu Leu Cys Ala Arg Ser Leu
305                 310                 315                 320

ctt cga ggc ttc ctg ctg cag aac gag ttt gtg ggg ttc atg tgg cgg   1008
Leu Arg Gly Phe Leu Leu Gln Asn Glu Phe Val Gly Phe Met Trp Arg
                325                 330                 335

cag cgg gga cgg gtc atc agc ctg tgg gag cgg ctg gaa ttt gtc aat   1056
Gln Arg Gly Arg Val Ile Ser Leu Trp Glu Arg Leu Glu Phe Val Asn
            340                 345                 350

ggc tgg tac atc ctg ctc gtc acc agc gat gtg ctc acc atc tcg ggc   1104
Gly Trp Tyr Ile Leu Leu Val Thr Ser Asp Val Leu Thr Ile Ser Gly
        355                 360                 365

acc atc atg aag atc ggc atc gag gcc aag aac ttg gcg agc tac gac   1152
Thr Ile Met Lys Ile Gly Ile Glu Ala Lys Asn Leu Ala Ser Tyr Asp
    370                 375                 380

gtc tgc agc atc ctc ctg ggc acc tcg acg ctg ctg gtg tgg gtg ggc   1200
Val Cys Ser Ile Leu Leu Gly Thr Ser Thr Leu Leu Val Trp Val Gly
385                 390                 395                 400

gtg atc cgc tac ctg acc ttc ttc cac aac tac aat atc ctc atc gcc   1248
Val Ile Arg Tyr Leu Thr Phe Phe His Asn Tyr Asn Ile Leu Ile Ala
                405                 410                 415

aca ctg cgg gtg gcc ctg ccc agc gtc atg cgc ttc tgc tgc tgc gtg   1296
Thr Leu Arg Val Ala Leu Pro Ser Val Met Arg Phe Cys Cys Cys Val
            420                 425                 430

gct gtc atc tac ctg ggc tac tgc ttc tgt ggc tgg atc gtg ctg ggg   1344
Ala Val Ile Tyr Leu Gly Tyr Cys Phe Cys Gly Trp Ile Val Leu Gly
```

```
        435                 440                 445

ccc tat cat gtg aag ttc cgc tca ctc tcc atg gtg tct gag tgc ctg   1392
Pro Tyr His Val Lys Phe Arg Ser Leu Ser Met Val Ser Glu Cys Leu
    450                 455                 460

ttc tcg ctc atc aat ggg gac gac atg ttt gtg acg ttc gcc gcc atg   1440
Phe Ser Leu Ile Asn Gly Asp Asp Met Phe Val Thr Phe Ala Ala Met
465                 470                 475                 480

cag gcg cag cag ggc cgc agc agc ctg gtg tgg ctc ttc tcc cag ctc   1488
Gln Ala Gln Gln Gly Arg Ser Ser Leu Val Trp Leu Phe Ser Gln Leu
                485                 490                 495

tac ctt tac tcc ttc atc agc ctc ttc atc tac atg gtg ctc agc ctc   1536
Tyr Leu Tyr Ser Phe Ile Ser Leu Phe Ile Tyr Met Val Leu Ser Leu
            500                 505                 510

ttc atc gcg ctc atc acc ggc gcc tac gac acc atc aag cat ccc ggc   1584
Phe Ile Ala Leu Ile Thr Gly Ala Tyr Asp Thr Ile Lys His Pro Gly
        515                 520                 525

ggc gca ggc gca gag gag agc gag ctg cag gcc tac atc gca cag tgc   1632
Gly Ala Gly Ala Glu Glu Ser Glu Leu Gln Ala Tyr Ile Ala Gln Cys
    530                 535                 540

cag gac agc ccc acc tcc ggc aag ttc cgc cgc ggg agc ggc tcg gcc   1680
Gln Asp Ser Pro Thr Ser Gly Lys Phe Arg Arg Gly Ser Gly Ser Ala
545                 550                 555                 560

tgc agc ctt ctc tgc tgc tgc gga agg gac ccc tcg gag gag cat tcg   1728
Cys Ser Leu Leu Cys Cys Cys Gly Arg Asp Pro Ser Glu Glu His Ser
                565                 570                 575

ctg ctg gtg aat                                                  1740
Leu Leu Val Asn
            580
```

<210> 4
<211> 764
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Thr Ser Pro Ser Ser Ser Pro Val Phe Arg Leu Glu Thr Leu Asp
  1               5                  10                  15

Gly Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg Gly Lys Leu Asp Phe
            20                  25                  30

Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe Gln Gly Glu Asp Arg
        35                  40                  45

Lys Phe Ala Pro Gln Ile Arg Val Asn Leu Asn Tyr Arg Lys Gly Thr
    50                  55                  60

Gly Ala Ser Gln Pro Asp Pro Asn Arg Phe Asp Arg Asp Arg Leu Phe
65                  70                  75                  80

Asn Ala Val Ser Arg Gly Val Pro Glu Asp Leu Ala Gly Leu Pro Glu
```

```
                85                  90                  95

Tyr Leu Ser Lys Thr Ser Lys Tyr Leu Thr Asp Ser Glu Tyr Thr Glu
            100                 105                 110

Gly Ser Thr Gly Lys Thr Cys Leu Met Lys Ala Val Leu Asn Leu Lys
        115                 120                 125

Asp Gly Val Asn Ala Cys Ile Leu Pro Leu Leu Gln Ile Asp Arg Asp
    130                 135                 140

Ser Gly Asn Pro Gln Pro Leu Val Asn Ala Gln Cys Thr Asp Asp Tyr
145                 150                 155                 160

Tyr Arg Gly His Ser Ala Leu His Ile Ala Ile Glu Lys Arg Ser Leu
                165                 170                 175

Gln Cys Val Lys Leu Leu Val Glu Asn Gly Ala Asn Val His Ala Arg
            180                 185                 190

Ala Cys Gly Arg Phe Phe Gln Lys Gly Gln Gly Thr Cys Phe Tyr Phe
        195                 200                 205

Gly Glu Leu Pro Leu Ser Leu Ala Ala Cys Thr Lys Gln Trp Asp Val
    210                 215                 220

Val Ser Tyr Leu Leu Glu Asn Pro His Gln Pro Ala Ser Leu Gln Ala
225                 230                 235                 240

Thr Asp Ser Gln Gly Asn Thr Val Leu His Ala Leu Val Met Ile Ser
                245                 250                 255

Asp Asn Ser Ala Glu Asn Ile Ala Leu Val Thr Ser Met Tyr Asp Gly
            260                 265                 270

Leu Leu Gln Ala Gly Ala Arg Leu Cys Pro Thr Val Gln Leu Glu Asp
        275                 280                 285

Ile Arg Asn Leu Gln Asp Leu Thr Pro Leu Lys Leu Ala Ala Lys Glu
    290                 295                 300

Gly Lys Ile Glu Ile Phe Arg His Ile Leu Gln Arg Glu Phe Ser Gly
305                 310                 315                 320

Leu Ser His Leu Ser Arg Lys Phe Thr Glu Trp Cys Tyr Gly Pro Val
                325                 330                 335

Arg Val Ser Leu Tyr Asp Leu Ala Ser Val Asp Ser Cys Glu Glu Asn
            340                 345                 350

Ser Val Leu Glu Ile Ile Ala Phe His Cys Lys Ser Pro His Arg His
        355                 360                 365

Arg Met Val Val Leu Glu Pro Leu Asn Lys Leu Leu Gln Ala Lys Trp
    370                 375                 380

Asp Leu Leu Ile Pro Lys Phe Phe Leu Asn Phe Leu Cys Asn Leu Ile
385                 390                 395                 400

Tyr Met Phe Ile Phe Thr Ala Val Ala Tyr His Gln Pro Thr Leu Lys
                405                 410                 415
```

```
Lys Gln Ala Ala Pro His Leu Lys Ala Glu Val Gly Asn Ser Met Leu
            420                 425                 430

Leu Thr Gly His Ile Leu Ile Leu Leu Gly Gly Ile Tyr Leu Leu Val
        435                 440                 445

Gly Gln Leu Trp Tyr Phe Trp Arg Arg His Val Phe Ile Trp Ile Ser
    450                 455                 460

Phe Ile Asp Ser Tyr Phe Glu Ile Leu Phe Leu Phe Gln Ala Leu Leu
465                 470                 475                 480

Thr Val Val Ser Gln Val Leu Cys Phe Leu Ala Ile Glu Trp Tyr Leu
                485                 490                 495

Pro Leu Leu Val Ser Ala Leu Val Leu Gly Trp Leu Asn Leu Leu Tyr
            500                 505                 510

Tyr Thr Arg Gly Phe Gln His Thr Gly Ile Tyr Ser Val Met Ile Gln
        515                 520                 525

Lys Val Ile Leu Arg Asp Leu Leu Arg Phe Leu Leu Ile Tyr Leu Val
    530                 535                 540

Phe Leu Phe Gly Phe Ala Val Ala Leu Val Ser Leu Ser Gln Glu Ala
545                 550                 555                 560

Trp Arg Pro Glu Ala Pro Thr Gly Pro Asn Ala Thr Glu Ser Val Gln
                565                 570                 575

Pro Met Glu Gly Gln Glu Asp Glu Gly Asn Gly Ala Gln Tyr Arg Gly
            580                 585                 590

Ile Leu Glu Ala Ser Leu Glu Leu Phe Lys Phe Thr Ile Gly Met Gly
        595                 600                 605

Glu Leu Ala Phe Gln Glu Gln Leu His Phe Arg Gly Met Val Leu Leu
    610                 615                 620

Leu Leu Leu Ala Tyr Val Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met
625                 630                 635                 640

Leu Ile Ala Leu Met Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser
                645                 650                 655

Trp Ser Ile Trp Lys Leu Gln Lys Ala Ile Ser Val Leu Glu Met Glu
            660                 665                 670

Asn Gly Tyr Trp Trp Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu
        675                 680                 685

Thr Val Gly Thr Lys Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe
    690                 695                 700

Arg Val Glu Glu Val Asn Trp Ala Ser Trp Glu Gln Thr Leu Pro Thr
705                 710                 715                 720

Leu Cys Glu Asp Pro Ser Gly Ala Gly Val Pro Arg Thr Leu Glu Asn
                725                 730                 735
```

```
Pro Val Leu Ala Ser Pro Pro Lys Glu Asp Glu Asp Gly Ala Ser Glu
            740                 745                 750

Glu Asn Tyr Val Pro Val Gln Leu Leu Gln Ser Asn
        755                 760
```

<210> 5
<211> 764
<212> PRT
<213> Homo sapiens

<220>
<223> Xaa at position 667 may be any amino acid

<400> 5

```
Met Thr Ser Pro Ser Ser Ser Pro Val Phe Arg Leu Glu Thr Leu Asp
  1               5                  10                  15

Gly Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg Gly Lys Leu Asp Phe
             20                  25                  30

Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe Gln Gly Glu Asp Arg
         35                  40                  45

Lys Phe Ala Ser Gln Ile Arg Val Asn Leu Asn Tyr Arg Lys Gly Thr
     50                  55                  60

Gly Ala Ser Gln Pro Asp Pro Asn Arg Phe Asp Arg Asp Arg Leu Phe
 65                  70                  75                  80

Asn Val Val Ser Arg Gly Val Pro Glu Asp Leu Ala Gly Leu Pro Glu
                 85                  90                  95

Tyr Leu Ser Lys Thr Ser Lys Tyr Leu Thr Asp Ser Glu Tyr Thr Glu
            100                 105                 110

Gly Ser Thr Gly Lys Thr Cys Leu Met Lys Ala Val Leu Asn Leu Lys
        115                 120                 125

Asp Gly Val Asn Ala Cys Ile Leu Pro Leu Leu Gln Ile Asp Arg Asp
    130                 135                 140

Ser Gly Asn Pro Gln Pro Leu Val Asn Ala Gln Cys Thr Asp Asp Tyr
145                 150                 155                 160

Tyr Arg Gly His Ser Ala Leu His Ile Ala Ile Glu Lys Arg Ser Leu
                165                 170                 175

Gln Cys Val Lys Leu Leu Val Glu Asn Gly Ala Asn Val His Ala Arg
            180                 185                 190

Ala Cys Gly Arg Phe Phe Gln Asn Gly Gln Gly Thr Cys Phe Tyr Phe
        195                 200                 205

Gly Glu Leu Pro Leu Ser Leu Ala Ala Cys Thr Lys Gln Trp Asp Val
    210                 215                 220

Val Ser Tyr Leu Leu Glu Asn Pro His Gln Pro Ala Ser Leu Gln Ala
225                 230                 235                 240
```

```
Thr Asp Ser Gln Gly Asn Thr Val Leu His Ala Leu Val Met Ile Ser
                245                 250                 255

Asp Asn Ser Ala Glu Asn Ile Ala Leu Val Thr Ser Met Tyr Asp Gly
            260                 265                 270

Leu Leu Gln Ala Gly Ala Arg Leu Cys Pro Thr Val Gln Leu Glu Asp
        275                 280                 285

Ile Arg Asn Leu Gln Asp Leu Thr Pro Leu Lys Leu Ala Ala Lys Glu
    290                 295                 300

Gly Lys Ile Glu Ile Phe Arg His Ile Leu Gln Arg Glu Phe Ser Gly
305                 310                 315                 320

Leu Ser His Leu Ser Arg Lys Phe Thr Glu Trp Cys Tyr Gly Pro Val
                325                 330                 335

Arg Val Ser Leu Tyr Asp Leu Ala Ser Val Asp Ser Cys Glu Glu Asn
            340                 345                 350

Ser Val Leu Glu Ile Ile Ala Phe His Cys Lys Ser Pro His Arg His
        355                 360                 365

Arg Met Val Val Leu Glu Pro Leu Asn Lys Leu Leu Gln Ala Lys Trp
    370                 375                 380

Asp Leu Leu Ile Pro Lys Phe Phe Leu Asn Phe Leu Cys Asn Leu Ile
385                 390                 395                 400

Tyr Met Phe Ile Phe Thr Ala Val Ala Tyr His Gln Pro Thr Leu Lys
                405                 410                 415

Lys Gln Ala Ala Pro His Leu Lys Ala Glu Val Gly Asn Ser Met Leu
            420                 425                 430

Leu Thr Gly His Ile Leu Ile Leu Leu Gly Gly Ile Tyr Leu Leu Val
        435                 440                 445

Gly Gln Leu Trp Tyr Phe Trp Arg Arg His Leu Phe Ile Trp Ile Ser
    450                 455                 460

Tyr Thr Asp Ser Tyr Phe Glu Ile Leu Phe Leu Phe His Ser Leu Leu
465                 470                 475                 480

Thr Val Val Ser Leu Val Leu Cys Phe Leu Val Ile Glu Trp Tyr Leu
                485                 490                 495

Pro Leu Leu Val Ser Ala Leu Val Leu Gly Trp Leu Asn Leu Leu Tyr
            500                 505                 510

Tyr Thr Arg Gly Phe Gln His Thr Gly Ile Tyr Ser Val Met Ile Gln
        515                 520                 525

Lys Val Ile Leu Arg Asp Met Val Arg Phe Leu Val Ile Tyr Leu Val
    530                 535                 540

Phe Leu Phe Gly Phe Ala Val Ala Leu Val Ser Leu Ser Gln Glu Ala
545                 550                 555                 560

Trp Arg Pro Glu Ala Pro Thr Gly Pro Asn Ala Thr Glu Ser Val Gln
```

565 570 575

Pro Met Glu Gly Gln Glu Asp Glu Gly Asn Gly Ala Gln Tyr Arg Gly
580 585 590

Ile Leu Glu Ala Ser Leu Glu Leu Phe Lys Phe Thr Ile Gly Met Gly
595 600 605

Glu Leu Ala Phe Gln Glu Gln Leu His Phe Arg Gly Met Val Leu Leu
610 615 620

Leu Leu Leu Ala Tyr Val Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met
625 630 635 640

Leu Ile Ala Leu Met Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser
645 650 655

Trp Ser Ile Trp Lys Leu Gln Lys Ala Ile Xaa Val Leu Glu Met Glu
660 665 670

Asn Gly Tyr Trp Trp Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu
675 680 685

Thr Val Gly Thr Lys Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe
690 695 700

Arg Val Glu Glu Val Asn Trp Ala Ser Trp Glu Gln Thr Leu Pro Thr
705 710 715 720

Leu Cys Glu Asp Pro Ser Gly Ala Gly Val Pro Arg Thr Leu Glu Asn
725 730 735

Pro Val Leu Ala Ser Pro Pro Lys Glu Asp Glu Asp Gly Ala Ser Glu
740 745 750

Glu Asn Tyr Val Pro Val Gln Leu Leu Gln Ser Asn
755 760

<210> 6
<211> 966
<212> PRT
<213> Homo sapiens

<400> 6
Met Val Asn Ser Arg Arg Val Gln Pro Gln Pro Pro Gly Asp Ala Gly
1 5 10 15

Arg Ser Pro Ala Pro Arg Ala Ser Gly Pro Gly Arg Leu Val Ala Gly
20 25 30

Gly Ala Gly Leu Ala Val Pro Gly Gly Leu Gly Glu Gln Arg Gly Leu
35 40 45

Glu Ile Glu Met Glu Arg Ile Arg Gln Ala Ala Ala Arg Asp Pro Pro
50 55 60

Ala Gly Ala Ser Ala Ser Pro Ser Pro Pro Leu Ser Ser Cys Ser Arg
65 70 75 80

Gln Ala Trp Ser Arg Asp Asn Pro Gly Phe Glu Ala Glu Glu Asp Asp

```
                85                  90                  95

Asp Asp Asp Glu Val Glu Gly Glu Glu Gly Gly Met Val Val Glu Met
            100                 105                 110

Asp Val Glu Trp Arg Pro Gly Ser Arg Arg Ser Ala Ser Ser Ser Ala
        115                 120                 125

Val Ser Ser Val Gly Ala Arg Gly Arg Gly Leu Gly Ser Tyr Arg Gly
    130                 135                 140

Ala Ala His Leu Ser Gly Arg Arg Arg Arg Leu Glu Asp Gln Gly Ala
145                 150                 155                 160

Gln Cys Pro Ser Pro Ala Gly Gly Gly Asp Pro Leu His Arg His Leu
                165                 170                 175

Pro Leu Glu Gly Gln Pro Pro Arg Val Ala Trp Ala Glu Arg Leu Val
            180                 185                 190

Arg Gly Leu Arg Gly Leu Trp Gly Thr Arg Leu Met Glu Glu Ser Asn
        195                 200                 205

Ala Asn Arg Glu Lys Tyr Leu Lys Ser Val Leu Arg Glu Leu Val Thr
    210                 215                 220

Tyr Leu Phe Phe Leu Val Val Leu Cys Ile Leu Thr Tyr Gly Met Met
225                 230                 235                 240

Ser Ser Asn Val Tyr Tyr Tyr Thr Arg Thr Leu Ser Gln Leu Phe Ile
                245                 250                 255

Asp Thr Pro Val Ser Lys Thr Glu Lys Thr Asn Phe Lys Thr Leu Ser
            260                 265                 270

Ser Met Glu Asp Phe Trp Lys Phe Thr Glu Gly Ser Phe Leu Asp Gly
        275                 280                 285

Leu Tyr Trp Lys Ala Gln Thr Ser Asn His Thr Gln Ala Asp Asn Arg
    290                 295                 300

Ser Phe Ile Phe Tyr Glu Asn Leu Leu Leu Gly Val Pro Arg Leu Arg
305                 310                 315                 320

Gln Leu Arg Val Arg Asn Gly Ser Cys Ser Ile Pro Gln Asp Leu Arg
                325                 330                 335

Asp Glu Ile Lys Glu Cys Tyr Asp Val Tyr Ser Val Ser Ser Glu Asp
            340                 345                 350

Arg Ala Pro Phe Gly Pro Arg Asn Gly Thr Ala Trp Met Tyr Thr Ser
        355                 360                 365

Glu Lys Glu Leu Asn Gly Ser Ser His Trp Gly Ile Ile Ala Ser Tyr
    370                 375                 380

Ser Gly Ala Gly Tyr Tyr Leu Asp Leu Ser Arg Thr Arg Glu Glu Thr
385                 390                 395                 400

Ala Ala Gln Leu Ala Gly Leu Arg Arg Asn Phe Trp Leu Asp Arg Gly
                405                 410                 415
```

```
Thr Arg Ala Ala Phe Ile Asp Phe Ser Val Tyr Asn Ala Asn Ile Asn
            420                 425                 430

Leu Phe Cys Val Val Arg Leu Leu Ala Glu Phe Pro Ala Thr Gly Gly
        435                 440                 445

Val Val Pro Ser Trp Gln Phe Gln Pro Val Lys Leu Ile Arg Tyr Val
    450                 455                 460

Thr Ala Phe Asp Phe Phe Leu Ala Ala Cys Glu Ile Ile Phe Cys Phe
465                 470                 475                 480

Phe Ile Ile Tyr Tyr Val Val Glu Glu Ile Leu Glu Ile Arg Ile His
                485                 490                 495

Arg Leu Ser Tyr Phe Arg Ser Phe Trp Asn Cys Leu Asp Val Val Ile
            500                 505                 510

Val Val Leu Ser Val Val Ala Met Val Ile Asn Ile Tyr Arg Met Ser
        515                 520                 525

Asn Ala Glu Gly Leu Leu Gln Phe Leu Glu Asp Gln Asn Ser Phe Pro
    530                 535                 540

Asn Phe Glu His Val Ala Tyr Trp Gln Ile Gln Phe Asn Asn Ile Ser
545                 550                 555                 560

Ala Val Met Val Phe Leu Val Trp Ile Lys Leu Phe Lys Phe Ile Asn
                565                 570                 575

Phe Asn Arg Thr Met Ser Gln Leu Ser Thr Thr Met Ser Arg Cys Ala
            580                 585                 590

Lys Asp Leu Phe Gly Phe Thr Ile Met Phe Ser Ile Ile Phe Leu Ala
        595                 600                 605

Tyr Ala Gln Leu Ala Tyr Leu Val Phe Gly Thr Gln Val Asp Asp Phe
    610                 615                 620

Ser Thr Phe Gln Glu Cys Ile Phe Thr Gln Phe Arg Ile Ile Leu Gly
625                 630                 635                 640

Asp Ile Asn Phe Ala Glu Ile Glu Glu Ala Asn Arg Val Leu Gly Pro
                645                 650                 655

Leu Tyr Phe Thr Thr Phe Val Phe Phe Met Phe Phe Ile Leu Leu Asn
            660                 665                 670

Met Phe Leu Ala Ile Ile Asn Asp Ser Tyr Ser Glu Val Lys Ser Ser
        675                 680                 685

Val Ala Gln Gln Lys Ala Glu Met Glu Leu Ser Asp Leu Ile Arg Lys
    690                 695                 700

Gly Cys Gln Lys Ala Leu Val Lys Leu Lys Leu Lys Arg Asn Thr Val
705                 710                 715                 720

Asp Ala Ile Ser Glu Ser Leu Arg Gln Gly Gly Gly Lys Leu Asn Phe
                725                 730                 735
```

```
Asp Glu Leu Arg Gln Asp Leu Lys Gly Glu Gly His Thr Asp Ala Glu
            740                 745                 750

Ile Glu Ala Ile Phe Thr Lys Tyr Asp Gln Asp Gly Asp Gln Glu Leu
        755                 760                 765

Thr Glu Arg Glu His Gln Gln Met Arg Asp Asp Leu Glu Lys Glu Arg
    770                 775                 780

Glu Asp Leu Asp Leu Glu His Ser Ser Leu Pro Arg Pro Met Ser Ser
785                 790                 795                 800

Arg Ser Phe Pro Arg Ser Leu Asp Asp Ser Glu Glu Glu Asp Asp Glu
                805                 810                 815

Asp Ser Gly His Ser Ser Arg Arg Arg Gly Ser Ile Ser Ser Gly Val
            820                 825                 830

Ser Tyr Glu Glu Phe Gln Val Leu Val Arg Arg Val Asp Arg Met Glu
        835                 840                 845

His Ser Ile Gly Ser Ile Val Ser Lys Ile Asp Ala Val Ile Val Lys
    850                 855                 860

Leu Glu Ile Met Glu Arg Ala Lys Leu Lys Arg Arg Glu Val Leu Gly
865                 870                 875                 880

Arg Leu Leu Asp Gly Val Ala Glu Asp Ala Arg Leu Gly Arg Asp Ser
                885                 890                 895

Glu Ile His Arg Glu Gln Met Glu Arg Leu Val Arg Glu Glu Leu Glu
            900                 905                 910

Arg Trp Glu Ser Asp Asp Ala Ala Ser Gln Thr Gly His Gly Val Ser
        915                 920                 925

Thr Gln Val Gly Leu Gly Gly Gln Pro His Pro Arg Asn Pro Arg Pro
    930                 935                 940

Pro Ser Ser Gln Ser Ala Glu Gly Leu Glu Gly Gly Gly Gly Asn Gly
945                 950                 955                 960

Ser Ala Asn Val His Ala
                965
```

<210> 7
<211> 1533
<212> PRT
<213> Homo sapiens

<400> 7

```
Met Tyr Ile Arg Val Ser Tyr Asp Thr Lys Pro Asp Ser Leu Leu His
  1               5                  10                  15

Leu Met Val Lys Asp Trp Gln Leu Glu Leu Pro Lys Leu Leu Ile Ser
             20                  25                  30

Val His Gly Gly Leu Gln Asn Phe Glu Met Gln Pro Lys Leu Lys Gln
         35                  40                  45
```

```
Val Phe Gly Lys Gly Leu Ile Lys Ala Ala Met Thr Thr Gly Ala Trp
     50                  55                  60

Ile Phe Thr Gly Gly Val Ser Thr Gly Val Ile Ser His Val Gly Asp
 65                  70                  75                  80

Ala Leu Lys Asp His Ser Ser Lys Ser Arg Gly Arg Val Cys Ala Ile
                85                  90                  95

Gly Ile Ala Pro Trp Gly Ile Val Glu Asn Lys Glu Asp Leu Val Gly
            100                 105                 110

Lys Asp Val Thr Arg Val Tyr Gln Thr Met Ser Asn Pro Leu Ser Lys
        115                 120                 125

Leu Ser Val Leu Asn Asn Ser His Thr His Phe Ile Leu Ala Asp Asn
    130                 135                 140

Gly Thr Leu Gly Lys Tyr Gly Ala Glu Val Lys Leu Arg Arg Leu Leu
145                 150                 155                 160

Glu Lys His Ile Ser Leu Gln Lys Ile Asn Thr Arg Leu Gly Gln Gly
                165                 170                 175

Val Pro Leu Val Gly Leu Val Val Glu Gly Gly Pro Asn Val Val Ser
            180                 185                 190

Ile Val Leu Glu Tyr Leu Gln Glu Glu Pro Pro Ile Pro Val Val Ile
        195                 200                 205

Cys Asp Gly Ser Gly Arg Ala Ser Asp Ile Leu Ser Phe Ala His Lys
    210                 215                 220

Tyr Cys Glu Glu Gly Gly Ile Ile Asn Glu Ser Leu Arg Glu Gln Leu
225                 230                 235                 240

Leu Val Thr Ile Gln Lys Thr Phe Asn Tyr Asn Lys Ala Gln Ser His
                245                 250                 255

Gln Leu Phe Ala Ile Ile Met Glu Cys Met Lys Lys Lys Glu Leu Val
            260                 265                 270

Thr Val Phe Arg Met Gly Ser Glu Gly Gln Gln Asp Ile Glu Met Ala
        275                 280                 285

Ile Leu Thr Ala Leu Leu Lys Gly Thr Asn Val Ser Ala Pro Asp Gln
    290                 295                 300

Leu Ser Leu Ala Leu Ala Trp Asn Arg Val Asp Ile Ala Arg Ser Gln
305                 310                 315                 320

Ile Phe Val Phe Gly Pro His Trp Thr Pro Leu Gly Ser Leu Ala Pro
                325                 330                 335

Pro Thr Asp Ser Lys Ala Thr Glu Lys Glu Lys Lys Pro Pro Met Ala
            340                 345                 350

Thr Thr Lys Gly Gly Arg Gly Lys Gly Lys Gly Lys Lys Lys Gly Lys
        355                 360                 365

Val Lys Glu Glu Val Glu Glu Glu Thr Asp Pro Arg Lys Ile Glu Leu
```

370 375 380

Leu Asn Trp Val Asn Ala Leu Glu Gln Ala Met Leu Asp Ala Leu Val
385 390 395 400

Leu Asp Arg Val Asp Phe Val Lys Leu Leu Ile Glu Asn Gly Val Asn
405 410 415

Met Gln His Phe Leu Thr Ile Pro Arg Leu Glu Glu Leu Tyr Asn Thr
420 425 430

Arg Leu Gly Pro Pro Asn Thr Leu His Leu Leu Val Arg Asp Val Lys
435 440 445

Lys Ser Asn Leu Pro Pro Asp Tyr His Ile Ser Leu Ile Asp Ile Gly
450 455 460

Leu Val Leu Glu Tyr Leu Met Gly Gly Ala Tyr Arg Cys Asn Tyr Thr
465 470 475 480

Arg Lys Asn Phe Arg Thr Leu Tyr Asn Asn Leu Phe Gly Pro Lys Arg
485 490 495

Pro Lys Ala Leu Lys Leu Leu Gly Met Glu Asp Asp Glu Pro Pro Ala
500 505 510

Lys Gly Lys Lys Lys Lys Lys Lys Lys Lys Glu Glu Glu Ile Asp Ile
515 520 525

Asp Val Asp Asp Pro Ala Val Ser Arg Phe Gln Tyr Pro Phe His Glu
530 535 540

Leu Met Val Trp Ala Val Leu Met Lys Arg Gln Lys Met Ala Val Phe
545 550 555 560

Leu Trp Gln Arg Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys
565 570 575

Lys Leu Tyr Lys Ala Met Ala His Glu Ser Ser Glu Ser Asp Leu Val
580 585 590

Asp Asp Ile Ser Gln Asp Leu Asp Asn Asn Ser Lys Asp Phe Gly Gln
595 600 605

Leu Ala Leu Glu Leu Leu Asp Gln Ser Tyr Lys His Asp Glu Gln Ile
610 615 620

Ala Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr
625 630 635 640

Cys Leu Lys Leu Ala Val Ala Ala Lys His Arg Asp Phe Ile Ala His
645 650 655

Thr Cys Ser Gln Met Leu Leu Thr Asp Met Trp Met Gly Arg Leu Arg
660 665 670

Met Arg Lys Asn Pro Gly Leu Lys Val Ile Met Gly Ile Leu Leu Pro
675 680 685

Pro Thr Ile Leu Phe Leu Glu Phe Arg Thr Tyr Asp Asp Phe Ser Tyr
690 695 700

```
Gln Thr Ser Lys Glu Asn Glu Asp Gly Lys Glu Lys Glu Glu Glu Asn
705                 710                 715                 720

Thr Asp Ala Asn Ala Asp Ala Gly Ser Arg Lys Gly Asp Glu Glu Asn
                725                 730                 735

Glu His Lys Lys Gln Arg Ser Ile Pro Ile Gly Thr Lys Ile Cys Glu
            740                 745                 750

Phe Tyr Asn Ala Pro Ile Val Lys Phe Trp Phe Tyr Thr Ile Ser Tyr
        755                 760                 765

Leu Gly Tyr Leu Leu Leu Phe Asn Tyr Val Ile Leu Val Arg Met Asp
    770                 775                 780

Gly Trp Pro Ser Leu Gln Glu Trp Ile Val Ile Ser Tyr Ile Val Ser
785                 790                 795                 800

Leu Ala Leu Glu Lys Ile Arg Glu Ile Leu Met Ser Glu Pro Gly Lys
                805                 810                 815

Leu Ser Gln Lys Ile Lys Val Trp Leu Gln Glu Tyr Trp Asn Ile Thr
            820                 825                 830

Asp Leu Val Ala Ile Ser Thr Phe Met Ile Gly Ala Ile Leu Arg Leu
        835                 840                 845

Gln Asn Gln Pro Tyr Met Gly Tyr Gly Arg Val Ile Tyr Cys Val Asp
    850                 855                 860

Ile Ile Phe Trp Tyr Ile Arg Val Leu Asp Ile Phe Gly Val Asn Lys
865                 870                 875                 880

Tyr Leu Gly Pro Tyr Val Met Met Ile Gly Lys Met Met Ile Asp Met
                885                 890                 895

Leu Tyr Phe Val Val Ile Met Leu Val Val Leu Met Ser Phe Gly Val
            900                 905                 910

Ala Arg Gln Ala Ile Leu His Pro Glu Glu Lys Pro Ser Trp Lys Leu
        915                 920                 925

Ala Arg Asn Ile Phe Tyr Met Pro Tyr Trp Met Ile Tyr Gly Glu Val
    930                 935                 940

Phe Ala Asp Gln Ile Asp Leu Tyr Ala Met Glu Ile Asn Pro Pro Cys
945                 950                 955                 960

Gly Glu Asn Leu Tyr Asp Glu Glu Gly Lys Arg Leu Pro Pro Cys Ile
                965                 970                 975

Pro Gly Ala Trp Leu Thr Pro Ala Leu Met Ala Cys Tyr Leu Leu Val
            980                 985                 990

Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Val Phe Asn Asn Thr
        995                1000                1005

Phe Phe Glu Val Lys Ser Ile Ser Asn Gln Val Trp Lys Phe Gln Arg
   1010                1015                1020
```

```
Tyr Gln Leu Ile Met Thr Phe His Asp Arg Pro Val Leu Pro Pro Pro
1025                1030                1035                1040

Met Ile Ile Leu Ser His Ile Tyr Ile Ile Ile Met Arg Leu Ser Gly
                1045                1050                1055

Arg Cys Arg Lys Lys Arg Glu Gly Asp Gln Glu Glu Arg Asp Arg Gly
            1060                1065                1070

Leu Lys Leu Phe Leu Ser Asp Glu Glu Leu Lys Arg Leu His Glu Phe
        1075                1080                1085

Glu Glu Gln Cys Val Gln Glu His Phe Arg Glu Lys Glu Asp Glu Gln
    1090                1095                1100

Gln Ser Ser Ser Asp Glu Arg Ile Arg Val Thr Ser Glu Arg Val Glu
1105                1110                1115                1120

Asn Met Ser Met Arg Leu Glu Glu Ile Asn Glu Arg Glu Thr Phe Met
                1125                1130                1135

Lys Thr Ser Leu Gln Thr Val Asp Leu Arg Leu Ala Gln Leu Glu Glu
            1140                1145                1150

Leu Ser Asn Arg Met Val Asn Ala Leu Glu Asn Leu Ala Gly Ile Asp
        1155                1160                1165

Arg Ser Asp Leu Ile Gln Ala Arg Ser Arg Ala Ser Ser Glu Cys Glu
    1170                1175                1180

Ala Thr Tyr Leu Leu Arg Gln Ser Ser Ile Asn Ser Ala Asp Gly Tyr
1185                1190                1195                1200

Ser Leu Tyr Arg Tyr His Phe Asn Gly Glu Glu Leu Leu Phe Glu Asp
                1205                1210                1215

Thr Ser Leu Ser Thr Ser Pro Gly Thr Gly Val Arg Lys Lys Thr Cys
            1220                1225                1230

Ser Phe Arg Ile Lys Glu Glu Lys Asp Val Lys Thr His Leu Val Pro
        1235                1240                1245

Glu Cys Gln Asn Ser Leu His Leu Ser Leu Gly Thr Ser Thr Ser Ala
    1250                1255                1260

Thr Pro Asp Gly Ser His Leu Ala Val Asp Asp Leu Lys Asn Ala Glu
1265                1270                1275                1280

Glu Ser Lys Leu Gly Pro Asp Ile Gly Ile Ser Lys Glu Asp Asp Glu
                1285                1290                1295

Arg Gln Thr Asp Ser Lys Lys Glu Glu Thr Ile Ser Pro Ser Leu Asn
            1300                1305                1310

Lys Thr Asp Val Ile His Gly Gln Asp Lys Ser Asp Val Gln Asn Thr
        1315                1320                1325

Gln Leu Thr Val Glu Thr Thr Asn Ile Glu Gly Thr Ile Ser Tyr Pro
    1330                1335                1340

Leu Glu Glu Thr Lys Ile Thr Arg Tyr Phe Pro Asp Glu Thr Ile Asn
```

```
1345                1350                1355                1360

Ala Cys Lys Thr Met Lys Ser Arg Ser Phe Val Tyr Ser Arg Gly Arg
                1365                1370                1375

Lys Leu Val Gly Gly Val Asn Gln Asp Val Glu Tyr Ser Ser Ile Thr
            1380                1385                1390

Asp Gln Gln Leu Thr Thr Glu Trp Gln Cys Gln Val Gln Lys Ile Thr
        1395                1400                1405

Arg Ser His Ser Thr Asp Ile Pro Tyr Ile Val Ser Glu Ala Ala Val
    1410                1415                1420

Gln Ala Glu Gln Lys Glu Gln Phe Ala Asp Met Gln Asp Glu His His
1425                1430                1435                1440

Val Ala Glu Ala Ile Pro Arg Ile Pro Arg Leu Ser Leu Thr Ile Thr
                1445                1450                1455

Asp Arg Asn Gly Met Glu Asn Leu Leu Ser Val Lys Pro Asp Gln Thr
            1460                1465                1470

Leu Gly Phe Pro Ser Leu Arg Ser Lys Ser Leu His Gly His Pro Arg
        1475                1480                1485

Asn Val Lys Ser Ile Gln Gly Lys Leu Asp Arg Ser Gly His Ala Ser
    1490                1495                1500

Ser Val Ser Ser Leu Val Ile Val Ser Gly Met Thr Ala Glu Glu Lys
1505                1510                1515                1520

Lys Val Lys Lys Glu Lys Ala Ser Thr Glu Thr Glu Cys
                1525                1530
```

INTERNATIONAL SEARCH REPORT

Internati plication No
PCT/US 01/11442

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K16/18 C12Q1/68 G01N33/53
A61K38/17 A61K39/395 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 00 17222 A (HUAMN GENOME SCIENCES, INC.) 30 March 2000 (2000-03-30)	1-15
A	abstract page 2, line 7 - line 14 gene ID NO: 16 page 48, line 1 -page 51, line 26 page 92; table 1 page 94, line 1 -page 156, line 21 page 186, line 29 -page 210, line 23 page 311 -page 314; claims 1-16,18-21 SEQ ID NO: 26 figures PP14-15 SEQ ID NO: 63 figures PP44-46	16-30

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 August 2001	12/09/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Fuchs, U

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL, HEIDELBERG, FRG 'Online! 11 December 1998 (1998-12-11) NCI-CGAP: "tb24a12.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE: 2055262 3' similar to WP: R13A5.1 CE01370, mRNA sequence" Database accession no. AI307240 XP002176007 cited in the application	1-9
A	the whole document	10-15

X	DATABASE EMBL, HEIDELBERG, FRG 'Online! 12 July 1999 (1999-07-12) HILLIER, L. ET AL.: "au44h03.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone IMAGE: 2517653 3' similar to WP: R13A5.1 CE01370, mRNA sequence" Database accession no. AI816064 XP002176008 cited in the application	1-9
A	the whole document	10-15
X	-& DATABASE EMBL, HEIDELBERG, FRG 'Online! 12 July 1999 (1999-07-12) HILLIER, L. ET AL.: "au44h03.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone IMAGE: 2517653 5' similar to WP: R13A5.1 CE01370, mRNA sequence" Database accession no. AI815981 XP002176009	1-9
A	the whole document	10-15

P,X	WO 01 12662 A (INCYTE GENOMICS, INC.) 22 February 2001 (2001-02-22) abstract page 5, line 9 -page 10, line 16 page 23, line 32 -page 58, line 24 clone ID NO: 977658 page 74; table 1 SEQ ID NO: 13 page 81; table 2 SEQ ID NO: 50 page 91; table 3 page 107 -page 111; claims 1-27 SEQ ID NO: 13 figures PP12-13 SEQ ID NO: 50 figures PP44-45	1-15

	-/--	

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	SUN, M. ET AL.: "Mucolipidosis type IV is caused by mutations in a gene encoding a novel transient receptor potential channel" HUMAN MOLECULAR GENETICS, vol. 9, no. 17, 12 October 2000 (2000-10-12), pages 2471-2478, XP002176006 gene MLIV abstract page 2474, column 2, line 22 -page 2476, column 1, line 27 page 2477, column 1, line 30 - line 46	1,3,7,13
P,X	-& DATABASE EMBL, HEIDELBERG, FRG 'Online! 1 March 2001 (2001-03-01) SUN, M. ET AL.: "CDNA: FLJ22449 FIS, CLONE HRC09609 (MUCOLIPIN) (MUCOLIPIDOSIS TYPE IV PROTEIN) (MUCOLIPIN 1)" retrieved from HOMO SAPIENS Database accession no. Q9GZU1 XP002176010 the whole document	1,3,7

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Internati plication No
PCT/US 01/11442

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0017222 A	30-03-2000	AU 5928899 A	10-04-2000
		EP 1115735 A	18-07-2001
WO 0112662 A	22-02-2001	AU 6906800 A	13-03-2001